Author:
Ye, Wenting

Title:
Bilingual advantages in cognitive control: inconsistent results and large individual variability within groups via Mousetracker

Bilingual advantages in cognitive control: inconsistent results and large individual variability within groups via Mousetracker

Wenting Ye

School of Experimental Psychology

University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Master of Science by Research in the Faculty of Life Sciences

September 2018

15704

Abstract

Bilingual advantage (BA) in executive functions has been controversially discussed for decades, with complex and inconsistent empirical findings. This study conducted three classic tasks (i.e. Flanker, Simon and Spatial Stroop tasks) measuring executive functions with English monolingual, Chinese monolingual and Chinese-English bilingual young adults. Unlike traditional button-press experiments, participants in the current study provided their responses via mouse movement on MouseTracker. This novel technique is considered more sensitive than traditional measure to participants' decision/action dynamics which might be obscured by button-press experiments. However, the current results did not replicate our previous findings and even reported negative effects (i.e. monolinguals outperformed bilinguals), which motivated us to analyse the individual variability of the raw data more deeply. Large individual variability has been found within groups and within individuals, illustrating various response characteristics of individuals. Such large variability was extremely unexpected, which warrants further research. In conclusion, the current results would not support the BA claim, and standard analysis of mouse movements might obscure large individual variability. Further studies should be cautious of asserting BA before examining individual variability.

Acknowledgments

This project has taken me one year to accomplish. I would like to thank my supervisor Markus Damian, who has always been there to support me. This is our second project on this field and we definitely have learned more in the last year. I would also like to thank Prof. Qingqing Qu, who helped me collect data back in China, and Dr. Chris Kent who provided precious comments.

My friends in Bristol were always willing to listen to me when I was stressed or overwhelmed throughout the entire project. They also took time to proofread my draft of the thesis, which makes me feel safe and supported.

Finally, I would like to thank my parents who are based in China. They encouraged me when I was down, and afforded my living expenses, allowing me to concentrate on my study.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Table of Contents

List of Figures and Tables

Chapter 1: Introduction to Bilingual Advantages and Literature Review

1.1 General introduction and aim of the current project

Speaking more than one language provides bilinguals with a range of benefits. Apart from considerable economic, societal and cultural benefits of bilingualism, it has been discussed for a decade that bilingualism can lead to non-verbal advantages in cognitive control. The claim is that since bilinguals constantly engage in managing, switching or suppressing languages, such practice might convey their linguistic advantage to a broader advantage in non-linguistic tasks, e.g. executive function tasks as measures of cognitive control (Bialystok, 2009; Kroll & Bialystok, 2013). One popular account (Inhibitory Control model; Green, 1998) has emerged that bilinguals have a constant need to inhibit their nontarget language, so that a bilingual advantage (BA) would particularly emerge in conditions that involve conflict resolving. However, this claim has been challenged by inconsistent findings (e.g. Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2015, 2016). With the limits of this inhibitory control model, another account of bilingual advantage has developed, which focuses on attentional control (or monitoring ability) (e.g. Zhou & Krott, 2016a, 2016b). Attentional control is the ability to selectively focus on targets on a given activity and it has been argued that enhanced attentional control may result in 'global RT advantage' (i.e. overall faster responses across all conditions) in bilinguals (Costa, Hernández, & Sebastián-Gallés, 2008). Generally, either of the two accounts would not cover all existing positive findings for bilingual advantages in cognitive control. Additionally, the existence of bilingual advantage remains highly controversial because of a mix of positive and null effects.

The current project focused on temporal-dynamic information in participants on a series of executive function tasks, via the mousetracking paradigm. Mousetracking, a novel technique, could provide the mouse movement trajectories for every trial. Unlike traditional experiments via button presses, this paradigm not only records the response latency and accuracy, but also real-time processing during decision making. The main reason we adopted this technique was that the findings in this field are extremely inconsistent, with a large amount of positive and null effects (see a review by Paap & Greenberg, 2013). The traditional technique might not be able to explain these controversial results simply with response latency. Moreover, this is our second project using this paradigm. Our previous findings (Damian, Ye, Oh, & Yang, 2018) indicated substantial differences in mouse movement trajectories between English monolinguals and Chinese-English bilinguals, validating the mousetracking paradigm as a sensitive technique to detect potential bilingualism effects. The aim of the current study was to replicate our previous study with slight modification in experiment structure and with a third group of Chinese monolinguals. The following literature review will first demonstrate behavioural and neuroimaging evidence for bilingual advantages in cognitive control. Then the potential underlying mechanisms of bilingual advantages, two accounts of bilingual advantages we mentioned above will be criticised. How we moved to the novel technique will be discussed in Chapter 2, along with a detailed introduction to our previous study and the purpose of the current project.

1.2 Overview of bilingual advantage literature

It has been estimated that more than half of the people in the world speak at least two languages. Research on bilingual effects especially on cognitive functions has increased in recent decades. These studies were based on a view of experience-dependent neuroplasticity that experiences could affect one's brain development and cognitive systems (Pascual-Leone, Amedi, Fregni, & Merabet, 2005). For instance, London taxi drivers who need to navigate the city every day have been reported to have larger regions of the hippocampus that are responsible for spatial navigation (Maguire et al., 2000); formal education might contribute to slowing cognitive decline with ageing in adults (Stern, 2002; Kramer et al., 2004 cited in Bialystok, 2017). Therefore, language use, as the most intense and sustained experience that human beings engage in, should have a potential to change one's brain development and cognitive systems. Then bilingualism, the ability to speak more than one language, might lead to differences between monolinguals and bilinguals in these aspects.

Bilingualism effects

Two research aspects of bilingualism have been built up with growing evidence. First, there is evidence that bilinguals have weaker verbal skills in each language compared to their monolingual peers (Bialystok, Craik, & Luk, 2012). For example, in verbal fluency tasks, participants were required to generate as many words as they can in a minute that conform to a phonological or semantic cue. Bilinguals were consistently found to have worse performance especially in semantic fluency conditions (Bialystok et al., 2008; Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000), even though they could respond in their dominant language (Gollan & Ferreira, 2009). Bilinguals have been found to hold a smaller size of vocabulary in both languages than monolinguals (Oller and Eilers, 2002; Perani et al., 2003; Portocarrero, Burright and Donovick, 2007). Bilinguals also have been reported to encounter more interference in a lexical decision task (Ransdell & Fischler, 1987).

However, second, bilinguals have been claimed to have better cognitive control than monolinguals at all ages (Bialystok, Craik, & Luk, 2012). Bialystok, Craik, Klein and Viswanathan (2004) pointed out that bilingual advantages usually appear in the presence of misleading information and interference resolving, whereas little evidence supports bilingual advantages in tasks without misleading context. There is also evidence indicating potential benefits of bilingualism to delay cognitive ageing (Adesope, Lavin, Thompson, & Ungerleider, 2010; Bialystok et al., 2012), while cognitive control is negatively associated with ageing in monolinguals. In terms of bilingual disadvantages in verbal tasks, in this paper, 'bilingual advantage (BA)' will be referred to as the bilingual advantage in cognitive but not linguistic aspects.

Neuroimaging evidence

Apart from behavioural studies, others have focused on neuroimaging evidence in bilinguals. Neuroimaging studies compare monolinguals with bilinguals in their brain structure and function, investigating potential neural correlates for differences in cognitive control behaviours. There is evidence that higher density of grey matter in the left inferior parietal cortex (LIPG) has been reported in bilinguals, especially in early bilinguals (however, see a discussion in Klein et al., 2014), high-proficiency bilinguals (Mechelli et al., 2004 in Bialystok, 2009), and older bilinguals (Abutalebi et al., 2015). These results might provide evidence that the density of grey matter in the LIPG is associated with second language acquisition. Moreover, this region has also been found to be related to cognitive processes, such as working memory (Buchsbaum et al., 2005), indicating a potential neural foundation for bilingual advantage.

Other studies have focused more on grey matter density in frontal regions, as the frontal lobe has been identified its role of executive control for a long time. Abutalebi et al. (2012) concluded that both language control in bilinguals and general cognitive control processes engage a common neural system, namely the dorsal anterior cingulate cortex (ACC). They suggested that bilinguals use the ACC more efficiently in monitoring conflicts than monolinguals do. Olsen et al. (2015) also found that there is a decline in grey matter density with age in monolinguals but not bilinguals. Given these studies on grey matter volume, Bialystok interpreted that bilingualism could protect bilinguals from reduction of grey matter in frontal areas with ageing, and the sustained grey matter in the frontal lobe is associated with a superior cognitive control.

Luk, Bialystok, Craik and Grady (2011) found higher white matter integrity in the corpus callosum extending to the superior and inferior longitudinal fasciculi in older adults who were bilinguals compared with monolinguals. They proposed that this sustaining structural and functional connectivity might be a neural foundation of “cognitive research” in bilinguals. These neuroimaging findings also support the neuroplasticity view of bilingualism.

1.3 Underlying mechanism of bilingualism

Joint activation in bilinguals

How do bilinguals produce languages? Evidence has suggested that both languages in bilinguals would be activated, even though one of the languages is not relevant under current contexts. For example, when a Chinese-English bilingual is attending an English lecture, their Chinese translation may be implicitly activated even though it is irrelevant or not helpful in the current situation. This non-selective lexical access was well documented for speech comprehension (Dijkstra & van Heuven, 2010) and speech production in bilinguals (Bialystok, Craik & Luk, 2012). For example, Thierry and Wu (2004) asked participants (English monolingual controls and Chinese-English bilinguals) to decide whether pairs of English words were semantically related or not, with event-related potential (ERP) recordings simultaneously. Participants had no hint that half of the pairs included a repeated character when written in Chinese (e.g. English pair: Train-Ham, in Chinese: Huo Che-Huo Tui, creating semantically unrelated but form repeated condition). Their results showed that semantically unrelated but form (hidden repeated character) related condition, induced longer reaction times, higher error rates and larger N400 ERP shifts in bilinguals, indicating unconscious access to the Chinese translation of English words in Chinese-English bilinguals. Subsequent similar research by Thierry and Wu (2007) with a third Chinese monolingual group indicated that the amplitude of the N400 was smaller for the two Chinese groups (Chinese monolinguals and Chinese-English bilinguals). As the amplitude of the N400 was negatively associated with word similarity, not only semantic relatedness induced smaller N400, form relatedness could also lead to smaller N400. That is, even though bilinguals were not explicitly showed Chinese characters, the hidden repeated form could also trigger their access to Chinese implicitly. They also investigated joint activation on the basis of phonology rather than orthography of spoken language in their following research (Wu & Thierry, 2010) and found similar effects.

Studies on joint activation provide an understanding of linguistic and non-linguistic processing in bilinguals. First, joint activation in bilinguals indicates a language selection problem for bilinguals to resolve, which is not considered by monolinguals: bilinguals have to resolve competition between languages to achieve appropriate language selection, which requires extra attention resources (Bialystok et al., 2012). This might explain why bilinguals have disadvantages in linguistic abilities (e.g. verbal fluency and vocabulary size), because bilinguals might need to deal with an extra task with limited attention resources. However, the constant need for language selection might not be a complete disadvantage. The cognitive mechanism, which is responsible for language selection, is thought to be executive functions, and the consequences of this constant training will be discussed below.

Inhibitory Control Hypothesis

There is evidence that both languages in bilinguals are active even though one of them is not required in the current situation. Such joint activation indicates a constant need for bilinguals to language selection. To achieve appropriate language selection, one explanation comes from the Inhibitory Control Model (Green, 1998), which is based on the Supervisory Attentional System (SAS; Norman & Shallice, 1986). Based on this model, inhibition on unintended language is the consequence of language selection.

As bilingualism leads to long-lasting practice in inhibitory control, researchers started to study whether this linguistic training could convey advantages in non-verbal cognitive tasks which involve inhibitory control. Supporting evidence for this claim has been reported in a range of executive function tasks, with bilinguals outperforming monolinguals (see a review by Bialystok, 2017). For instance, studies have reported that bilinguals suffer less in incongruent conditions compared with monolinguals in children (Martin-Rhee & Bialystok, 2008), in adults and in older people (Bialystok, Craik, Klein, & Viswanathan, 2004), on tasks thought as indicators of inhibitory control (e.g. Simon task). Supporting evidence also found that bilinguals outperformed monolinguals in conditions that involved ignoring irrelevant information but not in conditions that did not (for example Bialystok, 2001). Hence, bilingual advantage emerges with the presence of interference.

However, some evidence has indicated that the inhibitory control hypothesis could not explain all the effects of bilingualism. Evidence against inhibitory control as the only explanation for effects of bilingualism, was from research that included preverbal toddlers and infants (Bialystok et al., 2010; Kovacs & Mehler, 2009). As it was believed that children under four years old had only preliminary control of comprehension and nearly no language production, any group differences (mono- vs. bilingual) detected could not attribute to experiences of inhibitory control on the nontarget language. For example, there was one study with preschool toddlers (Bialystok et al., 2010) that found significant differences between monolinguals and bilinguals in nonverbal cognitive tasks. Another evidence against the inhibitory control hypothesis is that Hilchey and Klein (2011) reviewed the literature and argued that bilinguals did not outperform monolinguals only in incongruent conditions but also in congruent conditions, suggesting a "global bilingual advantage", that is, bilinguals carry out overall faster response latencies and/or higher accuracy than monolinguals. This evidence might not argue against 'bilingual advantage' but could not be explained by the inhibitory control hypothesis.

These limits of the IC model are considered and Green and Abutalebi (2013) further developed the IC model to the Adaptive Control Hypothesis. In this more recent model, they proposed that bilinguals will recruit different control processes to achieve language selection under different language contexts. They identified three interactional contexts, namely single language, dual language and dense code-switching, which induce different demands on selection. Eight control processes have been distinguished (goal maintenance, conflict monitoring, interference suppression, salient cue detection, selective response inhibition, task disengagement, task engagement, opportunistic planning). For example, the single language context is the least demanding context, as individuals will only use one language in one environment (e.g. speaking English with colleagues but speaking the native language at home with family and friends). They predicted that under the single language context, individuals recruit goal maintenance and interference control to sustain their current goal and to avoid distractions from cross-language cues. Despite this framework provides understanding for different contexts, Bialystok (2017) pointed out that supporting evidence is "extremely preliminary". However, further study should consider more on how to determine bilinguals better, as bilinguals are affected by many factors, such as interactional contexts.

Executive function framework and attentional control

The inhibitory control hypothesis could not explain all results in this field, but it is still possible to account for partial results. Because of the limits of the inhibitory control hypothesis, the BA research has been extended to other components of executive function. A framework of executive function by Miyake et al. (2000) proposed that inhibition is one of three subcomponents of execution function, and three subcomponents (i.e. shifting, updating/monitoring, and inhibition) are moderately correlated with each other. Studies managed to measure other components by tasks that are thought to involve switching between goals and maintaining goals. For example, Prior a MacWhinney (2010), bilinguals have been reported to show smaller switching costs than their monolinguals, in a task-switching paradigm. The switching component is not the focus of this project, but it could be a good direction for future research.

Another aspect of executive function is illustrated as "monitoring", "updating" or "attentional control" (Engle, 2002; Miyake et al., 2000; Zhou & Krott, 2016a). Researchers probably defined this ability in a slightly different way but similarly, as the ability to selectively maintain focus on a given activity, which should account for the "global (usually response latencies) advantage". Miyake and colleagues (2000) defined monitoring ability as a relatively independent component of executive function. Engle and colleagues referred to "working memory as executive attention", to explain the monitoring system. They explained that working memory capacity is not a storage place but the ability to use limited working memory resources to constrain one's attention to maintain information on a given task. In a review on this literature by Bialystok (2017), she highlighted the differences between Miyake and colleagues' "monitoring component" and Engle's "executive attention", that Engle's executive attention is a continuous construct, but not a discrete component as proposed by Miyake and colleagues. Therefore, different intensity of bilingualism might lead to quantitatively (but not qualitatively) difference in executive attention. Engle also pointed out that working memory capacity is necessary for executive attention, which is not mentioned in Miyake et al. (2000). In general, these terms (i.e. monitoring ability, executive attention and attentional control) describe a similar construct.

Overall, the hypothesis based on attention indicates that bilingualism is likely to produce enhanced executive attention. However, it is not necessary that the two accounts (enhanced inhibitory control and enhanced executive attention) are mutually exclusive. In some cases, smaller congruency effects and overall faster response latencies could both exist (e.g. Bialystok et al., 2004; Bialystok et al., 2005; Costa et al., 2008; Martin-Rhee & Bialystok, 2008). The two accounts are developed to explain different bilingual advantages in non-verbal cognitive tasks.

Findings against the bilingual advantage

However, more recently, the bilingual advantage claim is severely challenged. Paap and colleagues (Paap & Greenberg, 2013; Paap et al., 2015, 2016) have listed a series of weaknesses in research on BA. In Paap and Greenberg (2013), they conducted 15 executive function tasks to investigate if there are any bilingual advantage in executive functions. Their results turned out that significant difference was found in only one task as a bilingual disadvantage, lending no support for bilingual advantage hypothesis. They also failed to replicate previous positive findings (Costa et al., 2008; Prior & MacWhinney, 2010) with task switching.

Paap and colleagues (2013) also indicated the lack of cross-task correlations. They argued that if executive function tasks are not intercorrelated, then the bilingual advantage in one task is likely to be task-specific. In their study, they found an even negative correlation between Flanker effects and Simon effects (two classic interference effects thought to be indicators of inhibitory control). Little or no correlation was found between other indicators of inhibitory control, switching or monitoring.

Apart from null effects and weak convergent validity, publication bias is another weakness in BA research. de Bruin, Treccani, and Della Sala (2015) found that there was a publication bias favouring findings with bilingual advantages over those with null effects or bilingual disadvantages. Moreover, published studies with bilingual advantages are mostly underpowered with a small sample size. De Bruin et al. (2015) did a meta-analysis and found an effect size of d=.3 of bilingual advantages in executive functions (including findings with and without positive effects). Then Bakker (2015, cited in Paap et al., 2016) calculated the estimated sample size of a well powered study, which is at least 139 participants per group. However, most studies would not meet this criterion. Overall, the research on BA remains controversial and inconclusive.

Chapter 2: Overview of Mousetracking Literature and Current Study

2.1 Introduction to Mousetracking paradigm

Since the literature of BA is mixed and inconsistent, there was research criticising the traditional analysis technique on this field. Zhou and Krott (2016a) proposed that the traditional analysis that uses mean response latencies as an index for a set of data would be inappropriate. They suggested that mean RT might be representative of a group performance, as response latencies are assumed to be normally distributed (Balota & Yap, 2011). However, human's response latencies have been known as not normally distributed but rightward skewed, with "extremely slow responses" as a tail. In Zhou and Krott's study (2016a), they found that aggressive data trimming of such 'outliers' might obscure a genuine bilingual advantage. They applied a technique called ex-Gaussian analysis on their data. The ex-Gaussian analysis is a distribution analysis of response latencies and response latencies are decomposed into two parameters: the normally distributed portion (μ) and an exponentially distributed portion(τ). The logic behind this analysis is: an enhanced inhibitory control would be indicated by smaller congruency effects in the Gaussian part (μ), while an enhanced executive attention (they termed it as 'attentional control') would be indicated by shorter tails (τ) in bilinguals. They compared Chinese-English bilinguals to English monolinguals over three interference tasks (i.e. Flanker, Simon and Spatial Stroop tasks). In the analysis of mean response latencies, they found smaller congruency effects (without data trimming) in bilinguals. However, then they found such smaller congruency effects were mainly due to the distribution tails (τ) rather than the Gaussian part (μ). Therefore, they interpreted these results as an enhanced attentional control in bilinguals compared to monolinguals rather than enhanced inhibitory control.

Zhou and Krott's study might provide an explanation for the inconsistent findings in this field, that severe data trimming might obscure a BA. Additionally, smaller congruency effects in bilinguals might not always mean enhanced inhibitory control, when if their response latencies are mainly affected by the distribution tails. According to their study, analysis of response latencies might not be the best way to study bilingual advantage, which motivates us to pursue another approach.

Traditionally, motor responses were considered functionally independent from cognitive processing, as motor responses were simply the outcomes of cognition. Therefore, motor responses would not reveal much about the dynamics of cognitive processing and researchers have to infer what has happened in the 'black box'. However, more recent theories disagree and suggest that even a simple movement could provide rich information about internal cognitive processes. Continuous and temporally dynamic motor responses might capture ongoing cognitive processing. There are evidence indicating that hand movements are yoked to mental dynamics, making it an ideal response for analysis. Discrete key presses are not ideal because they could not reveal the ongoing mental dynamics.

A framework of Embodied Choice proposed by Lepora and Pezzulo (2015), in which action performance is a part of the decision-making process, discussed how motion and cognition are interplayed. In their work, they discussed the limitations of two other models of decision making. First, most studies would believe in a serial process, in which decision and action are neurally separated and arranged linearly, so a decision is made first and then an action is generated. Based on this model, it is fine to adopt the button-press experiments. The second model they have discussed is the parallel model, in which decision is continuously revised and transferred to the action. This model also assumes that decision and action are independent, but decision and action can be deployed in parallel. For example, before the completion of a decision, uncertainty in the decision may reflect as an attraction to incorrect responses during a mouse movement tracking (i.e. deflection in the trajectories to the incorrect response). To be noted, these two models both consider decision as independent and action does not affect the decision-making process. However, the embodied choice model suggests that action is part of the decision-making process. They highlighted that the action is costly. For example, when a lion starts approaching one of two gazelles, it will be too costly to change its target, so the lion's action will bias its following decision making (i.e. continue chasing the same gazelle). That is, since action is costly, the decision maker should consider the trade-off between alternative responses after their initial choice. Therefore, before the completion of a decision/action, the brain has to consider both the decision component (i.e. making a correct decision) and the action component (i.e. minimising action costs) simultaneously. The embodied choice model is of importance to ecologically relevant decision making. In general, decision making and action performance are inseparable, so it is important to capture the dynamics between them rather than simply recording the outcomes of a decision/action interplay. Obviously, button-press experiments would not be appropriate in this case.

In our first study in BA, we adopted a technique of measuring the real-time changes in response of mouse movements (i.e. mousetracking paradigm) rather than button presses. A relatively novel technique, mousetracking paradigm, might at least reveal one's action dynamics during a cognitive task. This technique has been used in a range of studies, such as on prospective memory and decision making. In a standard mousetracking trial, to initiate a trial, participants should click on the "start" button at the bottom centre of the screen. Then a target displays on the screen and participants are required to respond to it by clicking on one of two response buttons at the left top or right top areas. For example, in a Spatial Stroop task, participants are asked to judge the direction of a single arrow (target), regardless of the position of it. Figure 1.1 is a demonstration of a Spatial Stroop task via mousetracking paradigm. The two trajectories represent average mouse movements for two conditions (Panel A for congruent condition: the target arrow points in the same direction as the position of the arrow, e.g. an arrow points to left and on the left side of screen; Panel B for incongruent condition: the target arrow points in the opposite direction as the position of itself, e.g. an arrow points to right and on the left side of the screen).

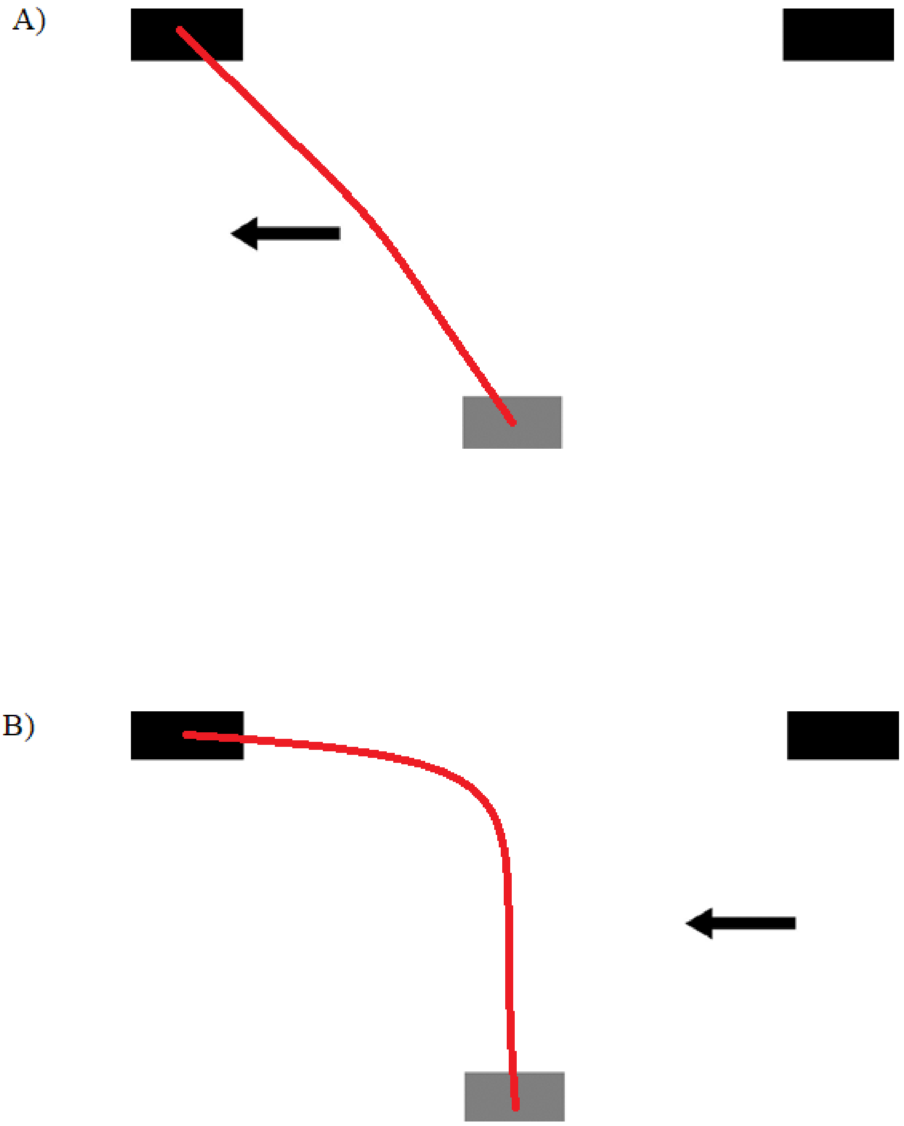


Figure 2.1. A demonstration of a Spatial Stroop task via mousetracking paradigm. Panel A represents a congruent trial. Panel B represents an incongruent trial. The red lines simulate a response.

Mouse movements from the "start" button to response button are continuously recorded as x- and y- coordinates pairs. Then a range of measures can be computed, including initiation time (i.e. duration between clicking on the start button and first mouse movement), response latency (duration between clicking on the start button and the response button), and measures related to curvature of the trajectory (see Figure 1.2): maximum deviation (MD; the largest perpendicular deviation from the idealised straight trajectory) and area under curve (AUC; the spatial area between the actual curve and the idealised straight line). A larger MD and AUC would represent a more deviated trajectory from the idealised trajectory (a straight trajectory between the 'start' button and response area). Then x-y coordinates pairs are illustrated visually. For a direct comparison, leftward trajectories would be horizontally flipped to the right side, hence all trajectories appear pointing rightward.

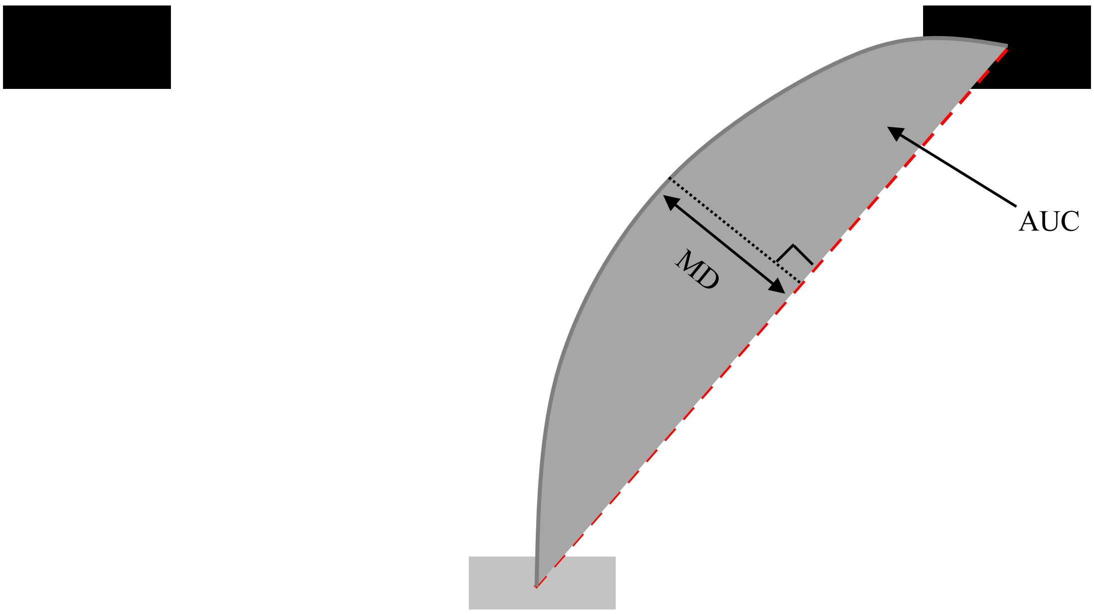


Figure 2.12. Measures of curvature of mouse trajectories: Maximum Deviation (MD) and Area Under Curve (AUC). The red line is the idealised trajectory and the grey line is a response.

In a review of bilingualism literature, only three previous studies investigated bilingual advantages via mousetracking paradigm (Incera & McLennan, 2016; 2018). For example, Incera and McLennan (2016) compared the performance of monolinguals to bilinguals in a Stroop task. In their task, participants were asked to judge the colour of the words but to ignore the text (e.g. blue in yellow font). They were instructed to make a response by clicking on response buttons on the screen (i.e. top left corner or top right corner). The results showed that bilinguals took longer initiation times and faster movements towards the correct response than monolinguals. They concluded that "bilinguals behave like experts" at conflict resolving, because bilinguals would withhold their response for a while and then make a more efficient response. In their subsequent research (Incera & McLennan, 2018), they investigated the effects of bilingualism and age on the Stroop and Flanker tasks. They found that younger age and higher-proficiency bilinguals were associated with reduced Stroop effect, but no such effects found in the Flanker task. Their first study inspired our first project on bilingual advantage (Damian, Ye, Oh, & Yang, 2018), in which we conducted similar tasks with monolinguals and bilinguals (discussed below). All three studies would agree that mousetracking is a more sensitive tool than traditional key presses experiment to participants' responses to EF tasks.

2.2 Previous study

In our first study, English monolinguals were compared to Chinese-English bilinguals through three EF tasks, namely the Flanker, Simon and Spatial Stroop task, via mousetracking paradigm. The task design was adapted from Zhou and Krott (2016a) but in a mousetracking version. All three tasks were used to measure the ability to suppress irrelevant information or distraction (i.e. inhibitory control). Response latency was analysed to compare with previous traditional key-press studies. According to the "expert-like behaviour" in bilinguals proposed by Incera and McLennan (2016), we also measured participants' initiation time and mouse trajectories. We predicted that if a BA emerged as bilinguals' enhanced inhibitory control, then congruency effects (e.g. Flanker effects) should be smaller for response latencies and/or for mouse trajectories (i.e. carrying out straighter mouse movement typically in incongruent trials); if a BA emerged as superior attentional control, then bilinguals should show an overall faster response latency and/or general more efficient trajectories on both congruent and incongruent trials. In terms of the initiation time, an early movement was predicted in bilinguals.

For the results, we found that monolinguals and bilinguals had comparable response latencies. This is possible that null effects emerge in response latencies, as response latencies might be not sensitive enough to detect a BA. The results of mouse trajectories might confirm this, as a BA emerged in mouse movements. Bilinguals tended to withhold their first movement for a longer time and then executed a more "efficient" response (e.g. Figure 2.1, results from Simon task). This figure illustrates the mouse trajectories we found: monolinguals and bilinguals showed similar trajectories in congruent trials but bilinguals' trajectory for incongruent trials was substantially more ideal than monolinguals' for incongruent trials. This was consistent with those findings obtaining null effects in button-press studies. Bilinguals also showed overall more "ideal" mouse movements compared to monolinguals, that is, group effects on MD and AUC. To summarise what we found, both groups showed similar response latencies, but their mouse trajectories showed different patterns.

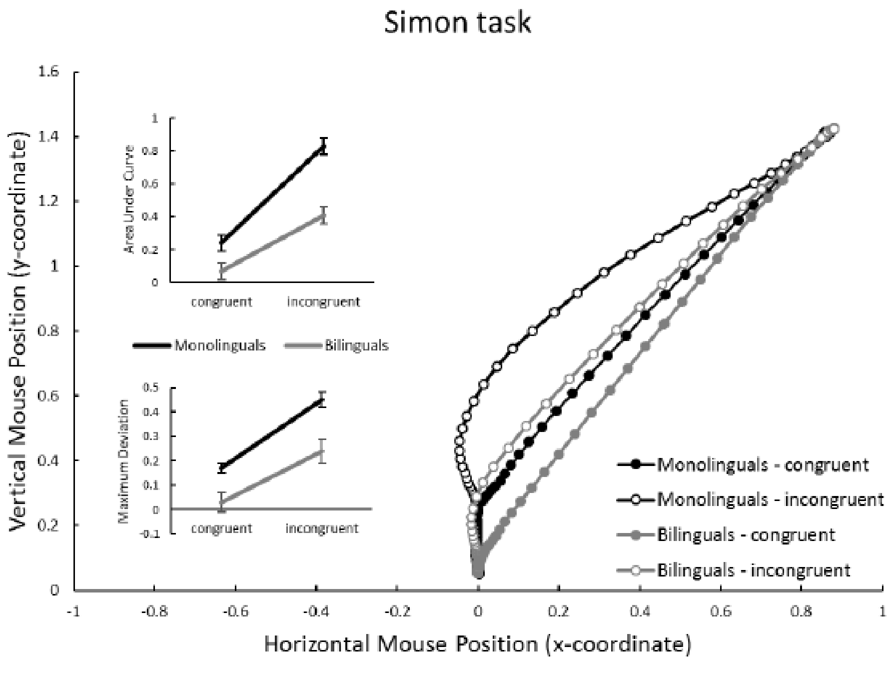


Figure 2.3. Simon task results from Damian, Ye, Oh, & Yang, 2018.

This study was interpreted to support the BA hypothesis that bilinguals enjoyed enhanced attentional control due to the group effects on mouse movement variables (i.e.

MD and AUC), even though not on response latencies. These results also indicated the limitation of button-press studies that might explain the highly controversial findings to some extent. However, as we recruited only one type of bilinguals (Chinese-English), it could be criticised that such group difference was due to cultural differences other than bilingualism. For example, bilinguals' later initiation could be because Chinese tended to avoid risk, while English monolinguals were less conservative. To rule out the potential cultural confound, it might be good to add a third Chinese monolingual group.

2.3 Current study

We aimed to replicate our first study adding a third Chinese monolingual group. This third-group idea was inspired by Yang, Yang and Lust (2010), in which they attempted to dissociate potential cultural effects from bilingualism. In their study, they compared Korean-English bilingual children to three monolingual groups (English and Korean monolinguals in the U.S.A and another Korean monolingual group in Korea) on the Attention Network Test (ANT). The ANT is similar to the Flanker task on the purpose of assessing their cognitive control. They found that bilinguals outperformed monolinguals on accuracy and response latencies, indicating a bilingualism effect on cognitive control. Although they also found cultural effects favouring Korean monolinguals from Korea, they suggested that bilingualism effects are more powerful than cultural effects at a cognitive level. However, their study was conducted with children. Studies on adults might illustrate whether this phenomenon persists over development.

The current study recruited Chinese-English bilinguals, English monolinguals from the UK and Chinese monolinguals from China. All participants would be young adults from university and come to the lab, going through the Flanker, Simon and Spatial Stroop tasks in one session. The settings of these tasks were nearly the same as our previous study, with a few modifications in experiment structure: 1) Given the fact that all participants initiated their responses before stimuli display (during fixation time) in our previous study, the fixation cross was removed in the current study. This would make the current study more similar to Incera and McLennan's one (2016) in procedure, as they displayed the stimuli right after participants have clicked on the "start" button; 2) Each task would contain more trials now, and for each task there would be three test blocks. Therefore, performance can be compared between adjacent blocks to see whether participants behave consistently over the task, in case there is any practice effect. 3) At the beginning of every test block, two buffer trials were added to help our participants pay attention to the task. All of these modifications were in order to normalise the task design.

Three groups (English mono- vs. Chinese-English bilinguals vs. Chinese mono-) were compared in terms of initiation times, response latencies, MD and AUC. We make the following predictions: First, incongruent trials should result in longer response latencies than congruent trials overall to confirm the validity of tasks. Second, bilinguals are predicted to show statistically later initiation times compared to the other two groups, according to what previous studies found as "expert behaviours"; Third, group effect or interaction between group and congruency on all dependent variables are examined to see if the current results support any bilingual advantage hypothesis. If bilinguals suffer less from congruency effects, it might be interpreted as enhanced inhibitory control; If bilinguals show overall more ideal trajectories than monolinguals, it will be interpreted as enhanced attentional control; Fourth, if the BA is genuine (rather than culture effects), then bilinguals are expected to outperform the other two groups, while the two monolingual groups have comparable performance on various measures (i.e. RT, initiation times, MD and AUC); if it is a group difference other than linguality, then the two Chinese groups should both outperform the English monolingual group.

Chapter 3 Experiment

3.1 Introduction

The Flanker, Simon and Spatial Stroop tasks were chosen to replicate previous studies. As mentioned above, bilinguals have been reported to show disadvantages in language proficiency and verbal fluency compared with monolinguals. Therefore, non-verbal tasks were often used to investigate their executive functions, as their responses would not be affected by their verbal abilities.

These three non-verbal tasks have been widely adopted as measures of executive functions in the bilingual literature. The reason to use three tasks rather than a single task is due to task impurity problem (Rabbitt, 1997). These tasks first were developed to measure inhibition only. However, researchers argued that these tasks do not measure ‘pure’ components and involve abilities such as attentional control or working memory capacity. It is nearly impossible to develop any ‘pure’ task, so Paap and Greenberg (2013) suggested that using multiple tasks (that all target one particular executive function component) to target their common cognitive control ability. In our case, the Spatial Stroop task and the Flanker task do not rely much on memory during the experiment, whereas the Simon task will require the participants to keep a stimulus-response association in mind all the time. Therefore, if a pattern consistently occurs across tasks, then it is unlikely that it is an effect of a specific task.

All these three tasks require the abilities to resolve conflicts, ignoring irrelevant information and maintaining task goals during the experiment. Hence, inhibitory control and attentional control are two components that these tasks are assumed to commonly capture. Congruency effects (Flanker effect, Simon effect or Stroop effect; slower response time) are expected generally across groups, indicating efforts on resolving conflicts in incongruent trials.

3.2 Method

3.2.1 Participants

Students recruited from the University of Bristol or Chinese universities took part in the experiment. Some English monolinguals took part in the experiment for obtaining course credits. Other English monolinguals and bilinguals were paid for their time. Chinese (quasi) monolinguals were volunteers from universities in Beijing recruited by our Chinese collaborator.

English monolinguals (N=79, age mean=21.32) all signed on the consent form and confirmed that they were native English speakers and "were not fluent in any other language". Chinese monolinguals (N=34, age mean=20.03) also confirmed that their native language was Chinese (Mandarin or Cantonese) and were not fluent in any other language. Chinese-English bilinguals (N=26, age mean=23.97) all confirmed that they were native speakers of Mandarin or Cantonese and were also fluent in English. Bilinguals were required to fill in a language history questionnaire which is adapted from Silverberg and Samuel (2004; cited in Zhou & Krott, 2016a) and used by our previous study, which asked for information about self-rated language proficiency, age of acquisition, and the switch pattern of languages (e.g. how often do you switch between English and Chinese; Appendix A). All participants had a normal or corrected-to-normal vision, aged over 18, and could comfortably use a mouse with the right hand.

Table 3.1 presents the demographic information of English mono-, Chinese (quasi) mono- and bilingual participants. To be noted, it is nearly impossible to find pure Chinese monolinguals from Chinese universities. All 'Chinese monolinguals' were exposed to English classes to some extent, whereas they were not immersed in an English-speaking country and not using English in daily life. In addition, most of these Chinese monolinguals participants did not consider English as one of their fluent languages and their self-rated proficiency in English (if available) is much lower than Chinese-English bilinguals. However, we have to admit that some might argue this group as ‘low-proficiency bilinguals’. Compared to the cohort of participants from our previous study, English monolinguals and Chinese-English bilinguals were all students from University of Bristol. Any student who had participated our previous study would not be allowed to participate the current study.

All participants signed a consent form at the beginning and at the end of the experiment to confirm that they were willing to provide their data for further analysis. All information would be kept confidential. Ethics were approved by the Faculty of Sciences at the University of Bristol.

Table 3.1 Demographics of mono- and bilingual speakers.

Variable	English monolingual	Chinese monolingual	Bilinguals
N	79	34	26
No. languages fluent	1.00	1.29	2.04
Mean age of English onset (yrs.)	Birth	N/A	7.44
Mean age of Chinese onset (yrs.)	N/A	Birth	0.69
Speak L2 fluently (self-rated proficiency%)	No	No	Yes (71%)
Speak L2 on a daily basis	No	No	25/26

3.2.2 General design and procedure

We adapted the procedure from the three tasks used in our last project, which adapted the procedure from Zhou and Krott (2016a) except that participants responded on MouseTracker rather than button presses. In the current work, the fixation cross was removed to make the experiment procedure as similar as Incera and McLennan’s study

(2016), in which they also had participants on Stroop task via mousetracking paradigm but displayed their stimuli immediately after their start click. The order of three tasks was rotated for all participants: Flanker-Simon-Stroop, Flanker-Stroop-Simon, Simon-Flanker-Stroop, Simon-Stroop-Flanker, Stroop-Flanker-Simon, or Stroop-Simon-Flanker. Trials in each task were randomised for every participant.

Stimuli were displayed on a 21-inch monitor with a 1920 x 1080 screen resolution and participants were sat about 60 cm in front of the computer screen. The experiment was operated on Windows 7 with a software called MouseTracker for data collection (Freeman and Ambady, 2010). In all the three tasks, participants were asked to start each trial by clicking on a grey button at the bottom centre of the screen, and to respond to a target by clicking on either the top left or the top right response button of the screen. Participants were required to make responses as quickly and accurately as they could.

Following the three tasks, monolinguals were instructed to complete a simple questionnaire about their demographics (e.g. age, sex) and bilinguals were to complete the language history questionnaire (adapted from Silverberg & Samuel, 2004 cited in Zhou & Krott, 2016a; see Appendix A). The whole session of experiment took approximately 30 minutes per subject.

3.2.3 Procedure

Flanker task

Materials and Design. The Flanker task adapted from the Erikson Flanker paradigm (Eriksen & Eriksen, 1974), then was used by Costa et al. (2009). Participants were displayed a string of five arrows. They were asked to judge the direction of the middle arrow and to ignore the four flanker arrows (two on each side) which either pointed in the same direction with the target (i.e. congruent condition: →→→→→ or ←←←←←), or in the opposite direction with the target (i.e. incongruent condition:←←→←← or →→←→→). The visual angle of the arrows was about .55

degrees, with a distance about .06 degrees between adjacent arrows.

Procedure. Participants initiated every trial by clicking on a grey button at the bottom centre of the screen. Once a trial was initiated, 50% of trials would have the stimuli appeared slightly above and 50% of those below the centre of the screen. The congruency of trials was 75/25, with 75% congruent trials and 25% incongruent trials. Costa et al. (2009) indicated that this would increase task difficulty, so participants were unable to predict easily the next trial. If the target points to the left, then participant would need to move their mouse to the top left of the screen and click on that response button; if the target points to the right, then participant moved to top right corner and click on the response button. A response must be made within 1,700ms or the stimulus would disappear. Participants were not given warning messages for slow initiations, slow latencies or errors.

Participants were instructed to respond as quickly and accurately as they could. The Flanker task consists of one practice block (with 32 trials) and three test blocks (32 trials for each block with two buffer trials at the beginning of each block). In total, there were 134 trials. Between adjacent blocks, participants were allowed to take a break. Stimuli were randomised every time and each participant had a different sequence of trials.

Simon task (Zhou & Krott, 2016a)

Materials and Design. Stimuli were 2.2 cm x 2.2 cm red and blue squares. Participants judged the colour of the stimuli and responded by clicking on the top left response button if the target was red, or on the top right response button if the target was blue. Targets were displayed either 7 cm left or 7 cm right to the centre of the screen, creating congruent trials (i.e. red square presented to the left or blue square to the left) and incongruent trials (i.e. red square to the right or blue square to the left).

Procedure. Participants initiated every trial by clicking on a grey button at the bottom centre of the screen. Participants were instructed to respond to left if there was a red square and respond to the right if a blue square, and to ignore the position of the square. Targets disappeared with a response or after 1,000 ms with no response. Participants were not given warning messages for slow initiations, slow latencies or errors.

There were two practice blocks to help participants associate the link between colour and response button. In the first practice block, the words "red" and "blue" were written in the left and right response buttons, respectively. Then in the second practice block, the words no longer appeared in the response buttons and participants had to respond based on their memory. There were 12 trials in each practice block.

Then three test blocks which were the same as the second practice block but more trials were followed. Each test block contains 24 trials and two buffers at the beginning. In total, participants would need to complete 102 trials in this task. Between adjacent block, Between adjacent block, participants were allowed to take a break. Stimuli were randomised every time and each participant had a different sequence of trials.

Spatial Stroop task

Materials and Design. Spatial Stroop task was adapted from the Simon task designed by Bialystok (2006), in which participants judged the direction of a single arrow. The arrow was 6.5 cm in length with a tail 0.5 cm in width. The widest point of the arrow was 1.5 cm. The arrow was displayed either 7 cm to the left or to the right of the centre, hence creating congruent trials (i.e. ← to the left, or → to the right) or incongruent trials (i.e. ← to the right or → to the left).

Procedure. Participants initiated every trial by clicking on a grey button at the bottom centre of the screen. Participants were instructed to respond to the direction of the arrow but to ignore the position of it. If the arrow points to left, then click on the left response button; if the arrow points to the right, then click on the right response button. Targets disappeared with a response or after 1,000 ms with no response. Participants were not given warning messages for slow initiations, slow latencies or errors.

This task consists of one practice block with 24 trials and three test blocks with 24 trials and 2 buffers each. In total, there were 102 trials in this task. Between adjacent block, Between adjacent block, participants were allowed to take a break. Stimuli were randomised every time and each participant had a different sequence of trials.

3.3 Results for standard analysis of mousetracking data

For each mouse movement, the software MouseTracker computed accuracy, initiation time (i.e. the duration between clicking on the start area and participant's first mouse movement), and response latencies (i.e. the time interval between target onset and participant clicking on the response button). Trajectories were rescaled into a standard MouseTracker coordinate space and all leftward responses were horizontally flipped to the right side for direct comparison. To quantify the extent to which participants were affected by irrelevant information, the MouseTracker computed two variables: Maximum Deviation (MD: the largest perpendicular deviation from the idealised straight trajectory) and Area under Curve (AUC: the spatial area between the actual curve and the idealised straight line). A larger MD and AUC indicates a participant makes a trajectory more deviated toward the incorrect response button.

Averages of initiation times, response latencies, maximum deviations (MD) and area under curve (AUC) for each of the three tasks, and for each group (English mono- vs. bilinguals vs. Chinese mono-) and congruency (congruent vs. incongruent), were statistically analysed by using a group x congruency two-way Analysis of Variance (ANOVA), computed separately for each task (Flanker, Simon, Spatial Stroop).

Overall accuracy was high (Flanker task: 99.7%, Simon task: 99.2%, Spatial Stroop task: 99.7%). This might not be unexpected, as participants had the opportunity to correct their responses (compared to key presses experiment), even if they initially made an incorrect movement. Accuracy then was not further analysed and trials that were incorrectly responded to were excluded from further analysis. Also, to make three tasks comparable, data from trials in Flanker task with response latencies longer than 1,000 ms were deleted, because the other two tasks had maximum response latencies as 1,000 ms.

3.3.1 Flanker

The statistical results of group x congruency two-way ANOVA were reported in Table 3.2. Four measures, including initiation time, response latency, maximum deviation (MD) and area under curve (AUC) were computed separately for each group (English monolinguals vs. bilinguals vs. Chinese monolinguals) and condition (congruent vs. incongruent). If any congruency x group interaction or group effect was found, then a follow-up ANOVA would be conducted to investigate the direction of effects.

Table 3.2. Initiation times, response latencies (ms), maximum deviation and area under curve, for the Flanker task, separately for each group (English monolingual vs. bilingual vs. Chinese monolingual) and condition (congruent vs. incongruent).

Participant group	**Initiation Times**	**Latencies**	**Max. Deviation**	**Area Under Curve**
English				
congruent	188	884	0.20	0.32
incongruent	187	1003	0.61	1.21
Effect	-1	119	0.41	0.89
Overall	188	944	0.41	0.77
Bilingual				
congruent	167	914	0.31	0.49
incongruent	167	1016	0.79	1.63
Effect	0	102	0.48	1.14
Overall	167	965	0.55	1.06
Chinese				
congruent	159	861	0.24	0.38
incongruent	158	1015	0.78	1.64
Effect	-1	154	0.54	1.26
Overall	159	938	0.51	1.01
ANOVA results				
Group	0.140	0.68	**<.001*** **	**<.001*** **
Congruency	0.900	**<.001*** **	**<.001*** **	**<.001*** **
Group x Congruency	0.980	**<.001*** **	**0.002** **	**<.001*** **
Follow-up analyses				
Group x Congruency				
English vs. Bilingual		0.11	0.06	**0.01***
Bilingual vs. Chinese		**<.001*** **	0.20	0.34
English vs. Chinese		**0.001*** **	**<.001**	**<.001*** **
Group				
English vs. Bilingual			**<.001*** **	**0.004** **
Bilingual vs. Chinese			0.32	0.59
English vs. Chinese			**0.005** **	**0.004** **

*Note. ***p < .001; **p < .01; *p < .05.*

Initiation and response times

Participants initiated their responses 171 ms on average after target onset. Figure 3.1 (panel A) provides a visualisation of the relation between initiation time, response latency and the associated response duration (response duration=response latency-initiation time) in the Flanker task. 0 indicates the onset of a trial.

From Figure 3.1 (panel A), the differences in initiation times among groups were subtle, which is confirmed by the ANOVA results. No main effect of group (p=.14) was found in initiation times, even if numerically the English monolingual group showed a later mean initiation times (English: mean = 188 ms; Bilingual: mean = 167 ms; Chinese: mean = 159 ms). Statistics also showed neither congruency effects ($F(1,141)=0.02$, $p=.90$, $\eta^2 <.001$) nor interaction of group x congruency ($F(2,141)=0.02$, $p=.98$, $\eta^2 <.001$) on initiation time. Therefore, three groups started their first movement at approximately the same time.

Response latencies from Table 3.2 also showed no effect of group ($F(2,141)=0.39$, $p=.68$, $\eta^2 =.005$). In contrast, there was a highly significant effect of congruency ($F(1,141)=783.69$, $p<.001$, $\eta^2 =.85$), which is visible from Figure 3.1 (panel A). There was also an interaction of group and congruency ($F(2,141)=9.64$, $p<.001$, $\eta^2 =.12$). In the follow-up analysis of the interaction, the Chinese monolingual group was found to have the largest congruency effects (Chinese vs. Bilingual: $F(1,63)=21.23$, $p<.001$, $\eta^2 =.25$; Chinese vs. English: $F(1,111)=11.40$, $p=.001$, $\eta^2 =.09$) compared with the other two groups on response latencies. This was reflected as a significantly larger Effect (=154ms; Effect=incongruent response latency – congruent response latency).

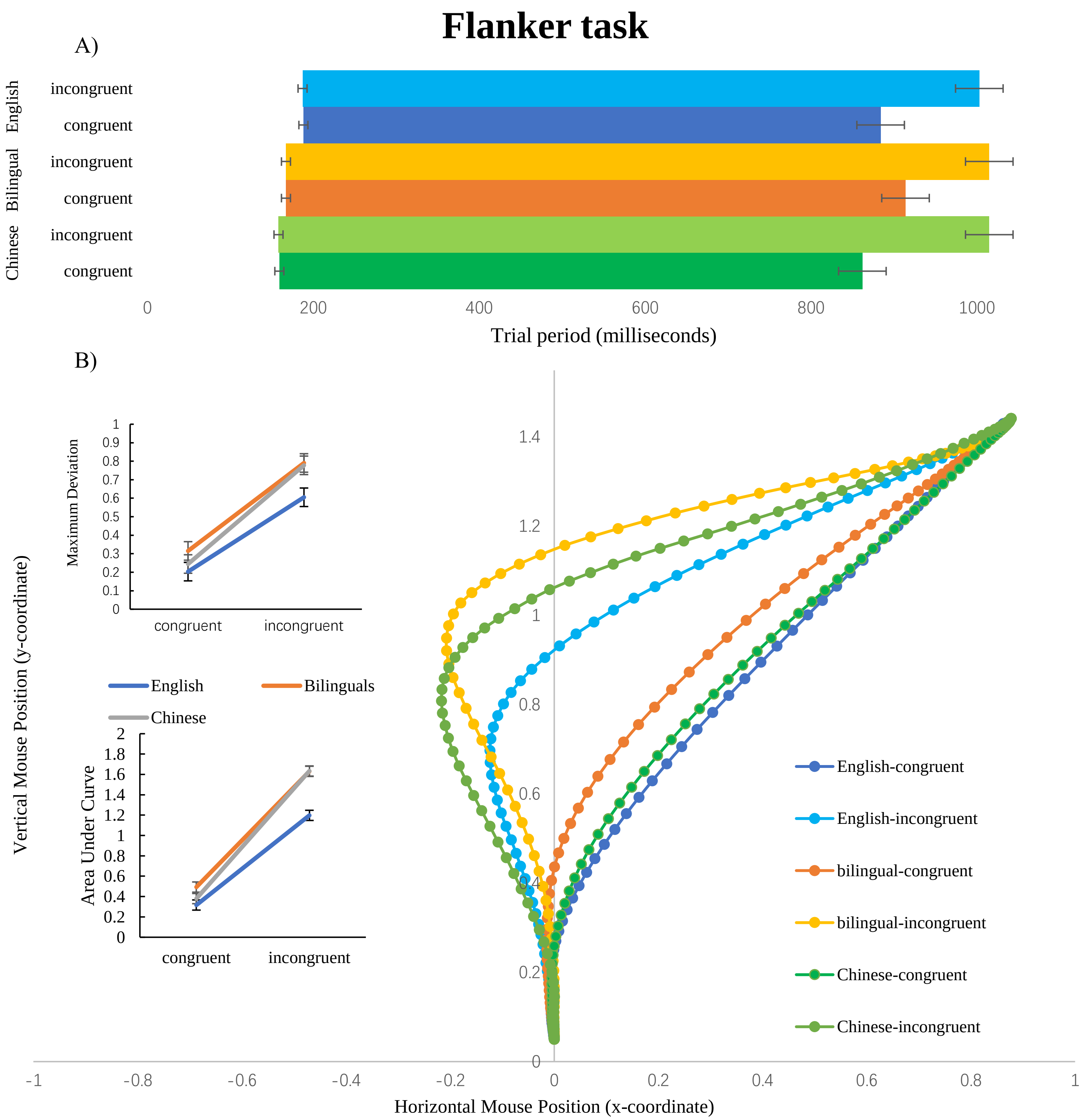


Figure 3.1 Panel A represents mouse movements for the Flanker task, separately for participant group and condition. 0 represents the onset of stimuli. Error bars represent standard error of the mean. Panel B represents mouse trajectories in x-y coordinate space for the Flanker task, separately for group and condition. Inset sub-panels are average Area under Curve and Maximum Deviation separately for group and congruency. Units for AUC are squared standard coordinates and units for MD are standard coordinates.

Mouse movement analysis

Maximum Deviation (MD) and Area under Curve (AUC) were computed to quantify the degree of deflection in mouse trajectories affected by the alternative response. Table 3.2 also reports an overview of MD and AUC results separately for each group (English mono- vs bilinguals vs Chinese mono-) and condition (congruent vs incongruent) for the Flanker task. Figure 3.1 (B) additionally showed average mouse trajectories separately for three groups and two conditions. As seen in Figure 3.1 (B), mouse trajectories were affected by congruency as expected, with incongruent responses illustrating a larger deflection towards the alternative response. This was reflected as a main effect of congruency on both MD ($F(1,141)=858.38$, $p< .001$, $\eta^2 =.86$) and AUC ($F(1,141)=609.99$, $p<.001$, $\eta^2 =.81$) (Table 3.2).

There was a main effect of group on MD ($F(2,141)=9.93$, $p<.001$, $\eta^2 =.12$) and AUC ($F(2,141)=8.55$, $p<.001$, $\eta^2 =.11$). In the follow-up analysis, English monolingual group showed the smallest MD (English vs. Chinese: $F(1,111)=8.37$, $p<.05$, $\eta^2 =.07$; English vs. Bilingual: $F(1,108)=16.68$, $p<.001$, $\eta^2 =.13$) and AUC compared to the other two groups (English vs. Chinese: $F(1,111)=8.66$, $p=.004$, $\eta^2 =.07$; English vs. Bilingual: $F(1,108)=13.54$, $p<.001$, $\eta^2 =.11$). An interaction on both MD ($F(2,141)=6.65$, $p=.002$, $\eta^2 =.09$) and AUC ($F(2,141)=7.70$, $p<.001$, $\eta^2 =.10$) was also found. In the follow-up analysis, English monolingual group has smaller congruency effects on MD compared to Chinese monolingual group ($F(1,111)=11.98$, $p<.001$, $\eta^2 =.10$); English monolingual group has the smallest congruency effects on AUC among groups (English vs. Chinese: $F(1,111)=12.95$, $p<.001$, $\eta^2 =.10$; English vs. Bilingual: $F(1,108)=6.30$, $p<.05$, $\eta^2 =.06$). Overall, the English monolingual group suffered the least from congruency effects and showed more “ideal” trajectories in both congruent and incongruent conditions. These are confirmed in the visualisation of mouse movement trajectories in Figure 3.1 (panel B), that English monolinguals overall showed more ideal trajectories, especially for incongruent conditions (the orange line; see also the inset panels in Figure for interaction).

3.3.2 Simon task

The statistical results of group x congruency two-way ANOVA were reported in Table 3.3. Four measures, including initiation time, response latency, maximum deviation (MD) and area under curve (AUC) were computed separately for each group (English monolinguals vs. bilinguals vs. Chinese monolinguals) and condition (congruent vs. incongruent). If any congruency x group interaction or group effect was found, then a follow-up analysis would be conducted to investigate the direction of effects.

Table 3.3. Initiation times, response latencies (ms), maximum deviation and area under curve, for the Simon task, separately for each group (English monolingual vs. bilingual vs. Chinese monolingual) and condition (congruent vs. incongruent).

Participant group	**Initiation Times**	**Response durations**	**Max. Deviation**	**Area Under Curve**
English				
congruent	154	595	0.10	0.16
incongruent	156	657	0.57	1.05
Effect	3	62	0.47	0.89
Overall	155	626	0.34	0.61
Bilingual				
congruent	130	632	0.16	0.22
incongruent	134	706	0.72	1.36
Effect	3	74	0.56	1.14
Overall	132	669	0.44	0.79
Chinese				
congruent	131	592	0.15	0.24
incongruent	130	657	0.62	1.17
Effect	-2	66	0.47	0.93
Overall	131	625	0.39	0.71
ANOVA results				
Group	0.10	0.11	**0.01***	0.05
Congruency	0.49	**<.001*****	**<.001*****	**<.001*****
Group x Congruency	0.61	0.33	0.08	0.09

Follow-up analyses	
Group	
English vs. Bilingual	**0.004****
Bilingual vs. Chinese	0.2
English vs. Chinese	0.15

*Note. ***p < .001; **p < .01; *p < .05.*

Initiation and response times

Participants initiated the movement remarkably earlier in the Simon task (average initiation time: 139 ms) than in the Flanker task (171ms). Table 3.3 showed no main effect of group, congruency and interaction on initiation times, which was showed in Figure 3.2 (panel A). The start points of each bar do not differ across conditions and groups. Again, the English monolingual group had a larger initiation time (English: mean = 155 ms vs. Bilingual: mean = 132 ms vs. Chinese: mean= 131 ms), even if it was not significant.

For response latencies, it was clear from Figure 3.2 (B) that congruency affected response latencies, as the orange bars are significantly longer than the blue bars. This was confirmed by the ANOVA results from Table 3.3 ($F(1,141)=542.10$, $p<.001$, $\eta^2 =.79$). In contrast, there was no group effect and no group x congruency interaction. That is, three groups had comparable response latencies, while incongruent trials took longer response latencies compared to congruent trials.

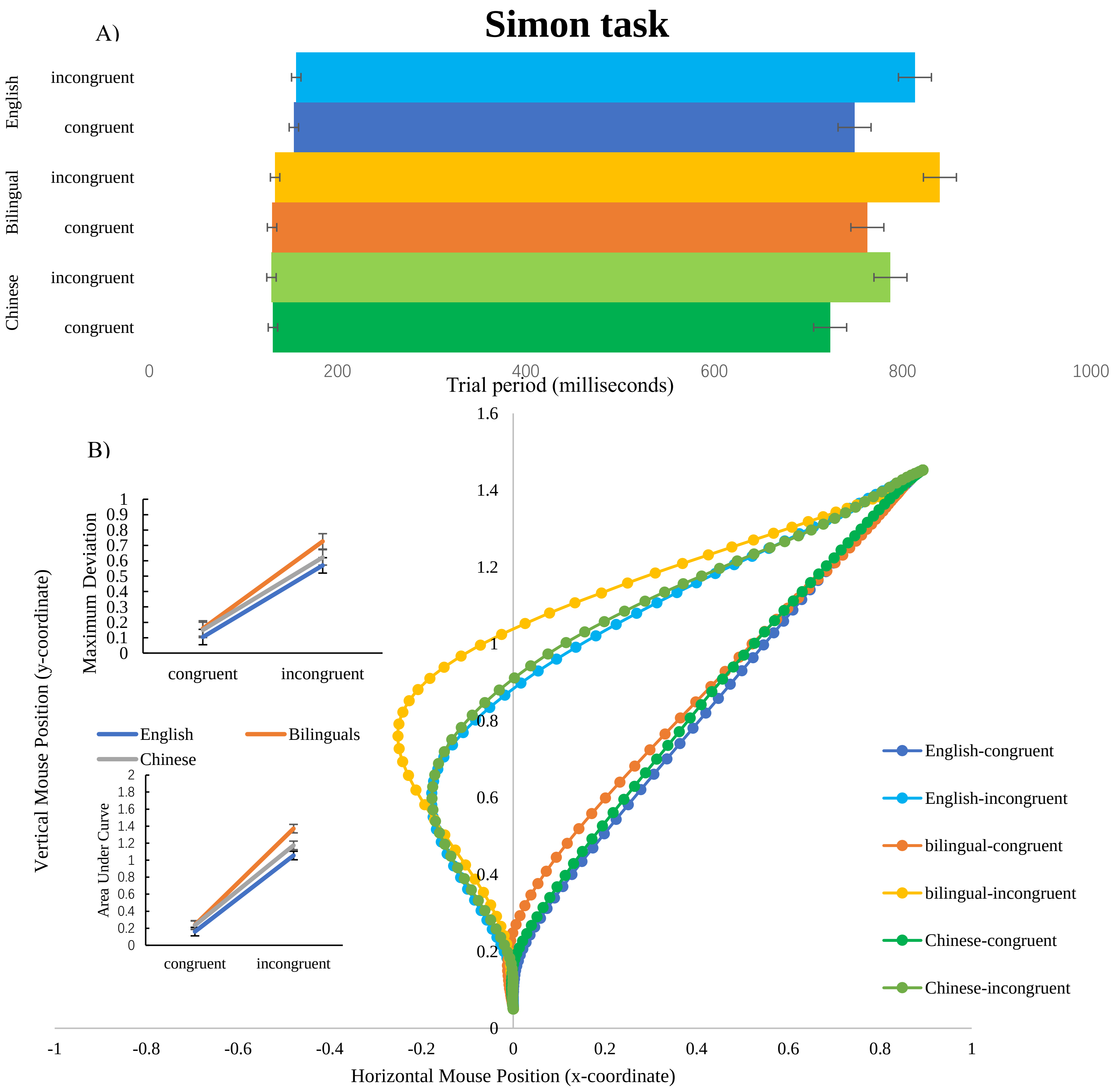


Figure 3.2. Panel A represents mouse movements for the Simon task, separately for participant group and condition. 0 represents the onset of stimuli. Panel B represents mouse trajectories in x-y coordinate space for the Simon task, separately for group and condition. Inset sub-panels are average Area under Curve and Maximum Deviation separately for group and congruency. Units for AUC are squared standard coordinates and units for MD are standard coordinates. Error bars represent standard error of the mean.

Mouse movement analysis

Mouse trajectories were affected by congruency (i.e. congruency effects on both MD and AUC), which was reflected in Figure 3.2 (panel B) as the large gap between congruent trajectories and incongruent trajectories (MD: $F(1,141)=690.64$, $p<.001$, $\eta^2=.83$; AUC: $F(1,141)=438.99$, $p<.001$, $\eta^2=.76$). Group effects only showed in MD ($F(2,141)=4.46$, $p=.01$, $\eta^2=.06$). In the follow-up analysis of the group effects, English monolinguals showed significantly smaller MD than bilinguals ($F(1,108)=8.53$, $p<.05$, $\eta^2=.07$), while the difference between English monolinguals and Chinese monolinguals or between bilinguals and Chinese monolinguals were not significant. As Figure 3.2 (panel B) indicated, the difference between English monolinguals and bilinguals was clear: English monolinguals carried out straighter response trajectories for both conditions than bilinguals; even though Chinese monolinguals were not significantly different from the other two groups from the statistics, the trajectories in Figure 3.2 (panel B) might show a more similar pattern between two monolingual groups.

3.3.3 Stroop

The statistical results of group x congruency two-way ANOVA were reported in Table 3.4. Four measures, including initiation time, response latency, maximum deviation (MD) and area under curve (AUC) were computed separately for each group (English monolinguals vs. bilinguals vs. Chinese monolinguals) and condition (congruent vs. incongruent). If any congruency x group interaction or group effect was found, then a follow-up analysis would be conducted to investigate the direction of effects.

Table 3.4. Initiation times, response latencies (ms), maximum deviation and area under curve, for the Spatial Stroop task, separately for each group (English monolingual vs. bilingual vs. Chinese monolingual) and condition (congruent vs. incongruent).

Participant group	**Initiation Times**	**Latencies**	**Max. Deviation**	**Area Under Curve**
English				
congruent	151	743	0.09	0.13
incongruent	152	811	0.61	1.14
Effect	1	69	0.52	1.01
Overall	151	777	0.35	0.64
Bilingual				
congruent	134	764	0.14	0.21
incongruent	127	844	0.79	1.57
Effect	-7	80	0.65	1.36
Overall	131	804	0.47	0.89
Chinese				
congruent	129	711	0.13	0.21
incongruent	131	777	0.70	1.36
Effect	2	66	0.57	1.15
Overall	130	744	0.42	0.79
ANOVA results				
Group	0.140	**0.02***	**0.006****	**0.005****
Congruency	0.400	**<.001*****	**<.001*****	**<.001*****
Group x Congruency	0.100	0.21	**0.03***	**0.01***
Follow-up analyses				
Group x Congruency interaction				
English vs. Bilingual			**0.006****	**0.003****
Bilingual vs. Chinese			0.08	0.11
English vs. Chinese			0.43	0.23
Group				
English vs. Bilingual		0.09	**0.002****	**0.002****
Bilingual vs. Chinese		**0.01***	0.24	0.27
English vs. Chinese		0.08	0.08	0.07

Note. ****p* < .001; ***p* < .01; **p* < .05.

Initiation and response times

Participants initiated their responses 137 ms on average after target onset, which was similar to the average initiation time in the Simon task (139ms) and earlier than that in the Flanker task (171ms). Like in the other two tasks, no main effect of group, congruency and interaction was found on initiation times in the Stroop task. This is showed in Figure 3.3 (panel A). Although the pattern of earlier mean initiation times for English group appeared again, their initiation times were not significantly smaller. In general, individuals did not differ in their initiation times.

In contrast, we found group effects on response latencies ($F(2,141)=4.08$, $p=.02$, $\eta^2=.05$) and the follow-up analysis showed a significant difference between bilinguals and Chinese monolinguals ($F(1,63)=6.79$, $p=.01$, $\eta^2=.10$). To be more specific, bilinguals showed significantly slower response latencies (mean=804 ms) than Chinese monolinguals (mean=744 ms) in both conditions (Table 3.4).

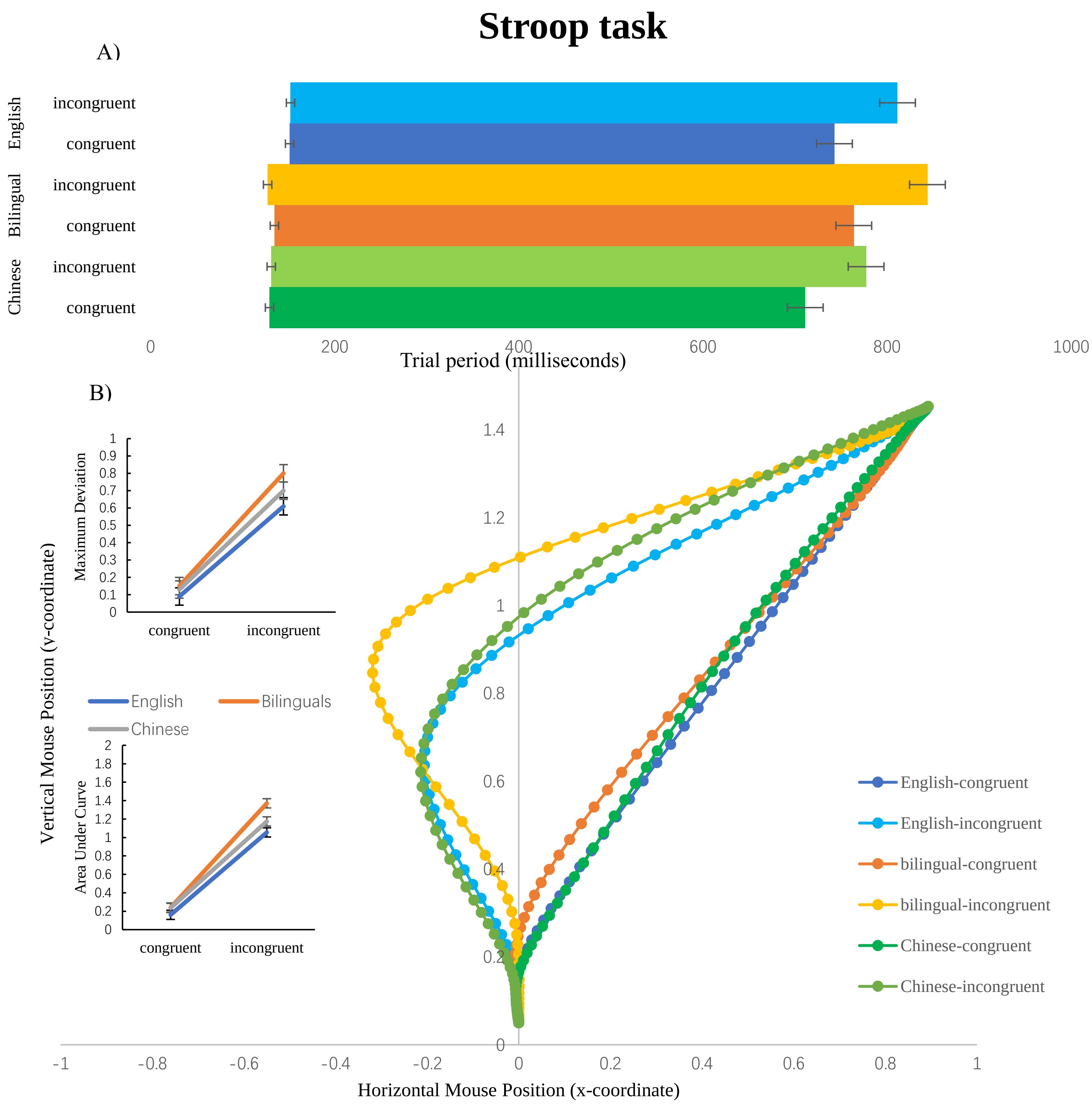


Figure 3.3. Panel A represents mouse movements for the Spatial Stroop task, separately for participant group and condition. 0 represents the onset of stimuli. Panel B represents mouse trajectories in x-y coordinate space for the Spatial Stroop task, separately for group and condition. Inset sub-panels are average Area under Curve and Maximum Deviation separately for group and congruency. Units for AUC are squared standard coordinates and units for MD are standard coordinates. Error bars are standard error of the mean.

Mouse movement analysis

The mouse trajectories showed a similar pattern as the Flanker results: besides congruency effects on MD (F(1,141)=877.25, p<.001, η^2=.86) and AUC (F(1,141)=533.25, p<.001, η^2=.79), a main effect of group and interaction were both found on MD (Group: F(2,141)=5.35, p=.006, η^2=.07; interaction: F(2,141)=3.78, p=.03, η^2=.05) and AUC (Group: F(2,141)=5.41, p=.005, η^2=.07; interaction: F(2,141)=4.72, p=.01, η^2=.06). In the follow-up analysis: English monolinguals showed a smaller MD (F(1,108)=9.93, p=.002, η^2=.08) and AUC (F(1,108)=10.39, p=.002, η^2=.09) compared to bilinguals; and English monolinguals showed significantly smaller congruency effects compared to bilinguals on both MD (F(1,108)=7.75, p=.006, η^2=.07) and AUC (F(1,108)=9.47, p=.003, η^2=.08). This is confirmed by the trajectories in Figure 3.3 (panel B) that, the English monolingual trajectories, both congruent one and the incongruent one, are more close to the "ideal" trajectories compared to the other two groups. Additionally, the difference between congruent and incongruent trajectories was significantly larger in bilinguals. Similar to the results from the Simon task, even though the trajectories from Chinese monolinguals are not significantly different from the other two groups, they are more close to trajectories from English monolinguals (Figure 3.3 panel B). Overall, English monolinguals carried out more "ideal" responses generally, especially on incongruent condition.

3.3.4 Summary

In all three tasks, there was no main effect of group, congruency and interaction on initiation times. However, the statistics showed a consistent pattern by group that English monolinguals started moving the mouse later than Chinese monolinguals later than bilinguals on average. For response latencies, bilinguals also showed a (not significantly) larger average response latencies, followed by English monolinguals and then Chinese monolinguals.

Given the mouse trajectories, in all three tasks, English monolinguals carried out a more ideal trajectory across conditions and tasks. Although in congruent trials, three groups showed similar trajectories over tasks, English mono- always had slightly more ideal trajectories and bilinguals had the most deviated ones. In incongruent condition, English mono- clearly had the most ideal movements and Chinese mono- showed very similar trajectories in the Simon and Stroop task, while bilinguals carried out considerably deviated trajectories over tasks. Interestingly, the results from the Simon task and the Spatial Stroop task were more similar than to the Flanker task, given their same range of initiation times, response latencies and the similar pattern of mouse movements across conditions.

Up to this point, the results we found were highly unexpected, as these findings do not only disagree with the BA claim that we found previously, but they would also not be expected by sceptics of the BA claim. More specifically, our previous work would have predicted earlier initiation times for mono- than for bilinguals and bilingual advantage on mouse trajectories; sceptics of BA claim would have predicted null effects on these measures. However, it would be odd to find 'monolingual advantage' on cognitive control. These findings certainly warrant a deeper analysis of the data, such as analysis of individual variability.

3.4 Results for individual variability

3.4.1 General individual average trajectories by group

In this section, a further analysis of the data was conducted on R (R Core Team, 2013) because of the unexpected "monolingual advantage" in three EF tasks. First, trajectories for per individual were mapped by group (Figure 3.4). Every trajectory is an average trajectory of all raw trajectories (by computing the means of x and y coordinates for all trials from all tasks) for one individual. These trajectories were not categorised by task or condition, but as an initial impression of an individual's mouse movement. From Figure 3.4, every grey line is an average trajectory for one individual and every colour (green, blue and red) line is an average of the individual average trajectories.

Surprisingly, huge individual variability was shown in all groups, with very straight lines and very deviated lines in each group, indicating that individuals could vary their responses within a group. With a glance at this figure, it is difficult to identify any clear 'cluster' of trajectories for each group. In addition, when comparing the three average group trajectories (left top panel in Figure 3.4), even though bilinguals have a less ideal trajectory, the differences among groups were not that large.

The individual variability which emerges in the figure is obscured in the average trajectories which are normally plotted in mousetracking experiments, such as our earlier figures (Figure 3.1-3.3) and previous findings (Damian, Ye, Oh, & Yang, 2018). This is also not a standard analysis of studies which had used mousetracking before, as we could not find any mousetracking study would have investigated the individual variability.

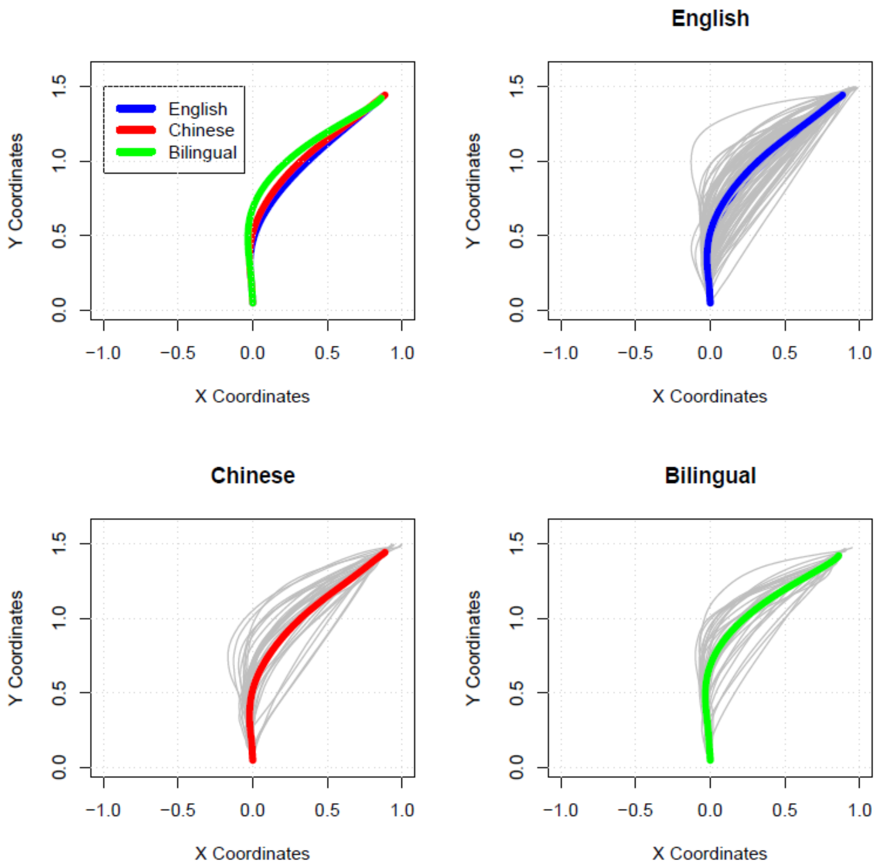
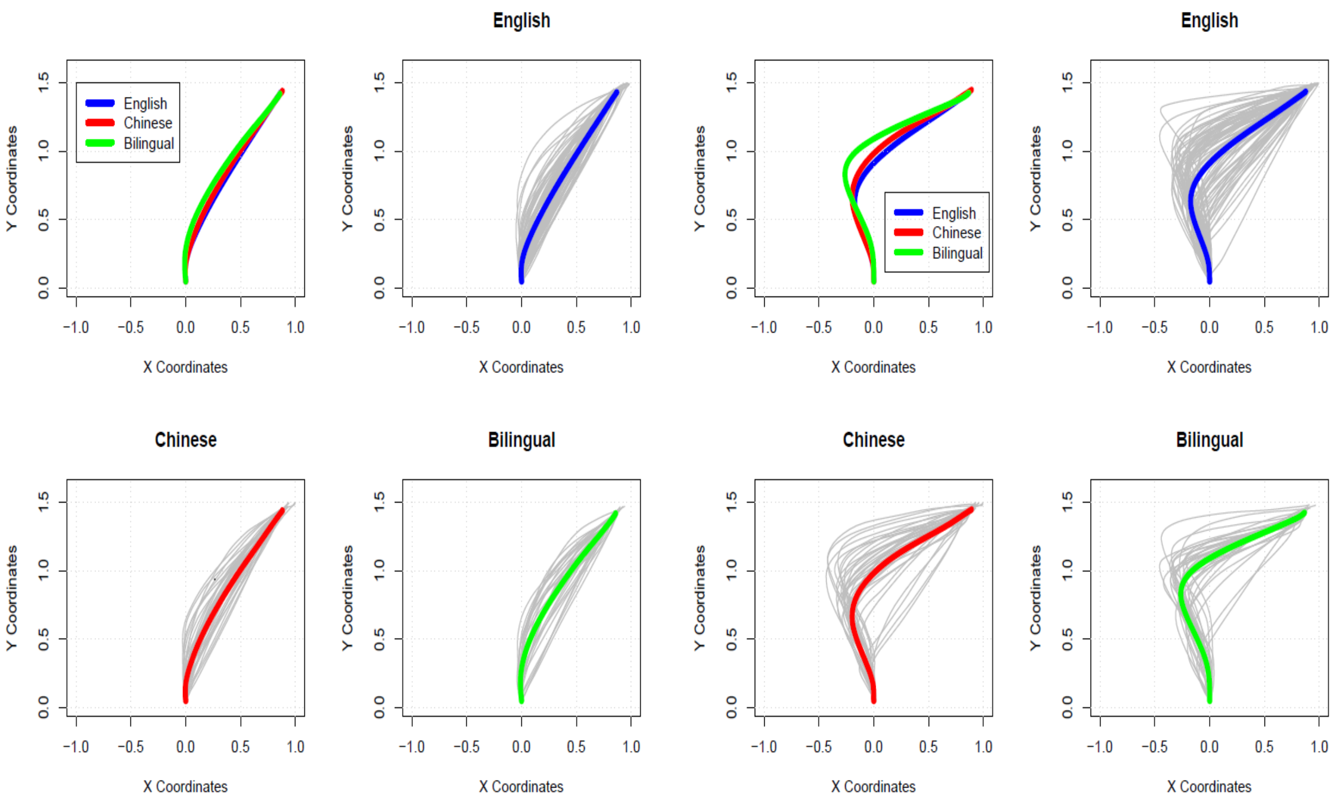


Figure 3.4. Individual average trajectories separately for each group. Every grey line represents an overall performance of one individual and the colour line is the average trajectories for each group (Blue for English monolinguals, red for Chinese monolinguals and Green for bilinguals)

3.4.2 Individual average trajectories by group for two conditions

The previous Figure 3.4 collapsed across conditions (congruent and incongruent) and tasks (Flanker, Simon and Spatial Stroop). However, from Figure 3.4 it is not clear whether this variability particularly emerges in which condition. This motivates splitting up the data by congruent/incongruent, as in Figure 3.5.



A) B)

Figure 3.5 Individual average trajectories separately for group and condition. Panel A represents the individual average trajectories for congruent condition and Panel B represents the individual average trajectories for incongruent condition. Every grey line represents an overall performance of one individual and the colour line is the average trajectories for each group (Blue for English monolinguals, red for Chinese monolinguals and green for bilinguals).

The individual average trajectories of congruent conditions and incongruent conditions were illustrated separately by group, with congruent trajectories as the baseline (Figure 3.5 Panel A). Every grey line in Panel A represents an average trajectory for one individual's congruent trials and every grey line in Panel B represents an average trajectory for one individual's incongruent trials. Every colour line is an average trajectory of all those individual trajectories in a group. From Figure 3.5, it is clear that participants did not differ much in congruent conditions by group, but showed great differences in incongruent conditions. More specifically, bilinguals showed the most deviated average trajectory among groups for the incongruent condition. These results are consistent with those results in our standard analysis of the current data. However, individuals showed large individual variability in incongruent trajectories. Some individuals carried out ideal and efficient trajectories, while others showed very deviated ones in incongruent condition.

3.4.3 Individual trial-by-trial trajectories by group

The individual average trajectories in Figure 3.5 might still obscure underlying variation, such as intra-variability. To dig into more of this variability, we plotted participants' raw trajectories from each group. On the left side of Figure 3.6, there are histograms of global MD for each group and subject, indicating that the majority of participants (in each group) have "average" MDs but that there are a few which have very small. This might suggest that some individuals might consistently made very efficient movements regardless of conditions. Then the right side of Figure 3.6 shows some examples (raw trajectories plotted from 16 participants from each group) from each group to show how exactly one individual would perform over the whole experiment. Every blue line is an individual's response to a congruent trial and every red line is a response to an incongruent trial. The black lines are the average of the blue lines or the average of the red lines, corresponding to one of those grey lines for congruent trial and incongruent trial, respectively. Normally, the average response of the congruent trials looks more ideal, whereas the average of the incongruent trials has deviated toward the incorrect responses. From Figure 3.6, we can also easily compare trajectories between two participants. For instance, participants 7 and 8 from the "English" group exhibit rather substantial differences in response characteristics. Participant 7 carried out very "efficient" responses in both congruent and incongruent trials, whereas participant 8 only showed "efficient" response in congruent trials. This confirms the inter-individual variability showed in Figure 3.5. We also plotted two participants on all three tasks separately in Figure 3.7 to see whether this inter-individual variability exists in a particular task. This figure demonstrated that such inter-individual variability is consistent across tasks, indicating substantial differences in response characteristics.

Interestingly, while participants' responses for congruent trials were clustered around the "ideal trajectory", their responses for incongruent trials were of high variation, suggesting high intra-individual variability.

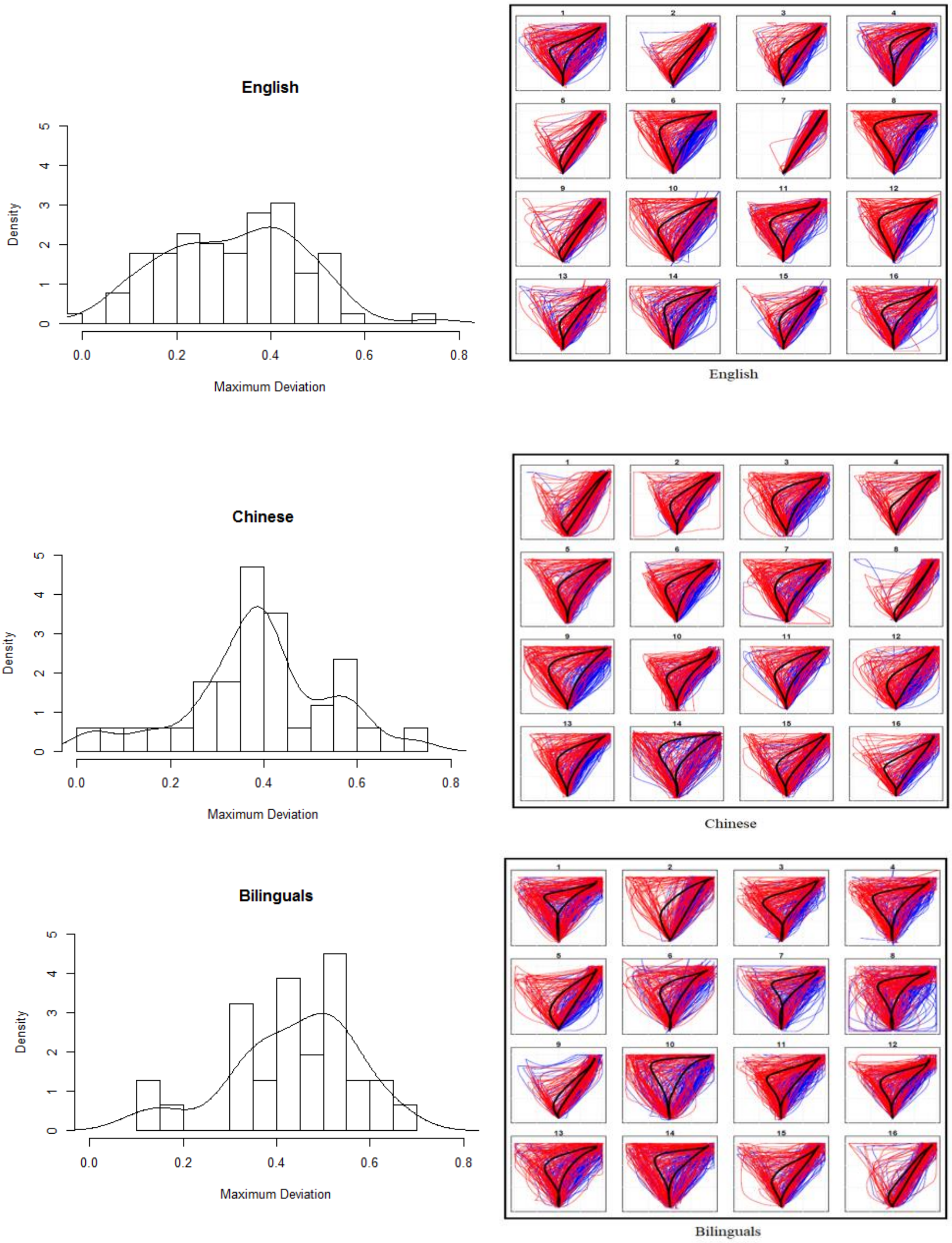


Figure 3.6. Left panels were histograms of global MDs for each group and subject. Right panels were 16 individuals' trajectories for each group (English monolinguals vs. Chinese monolinguals vs. bilinguals). Every blue line represents a congruent trial and every red line represents an incongruent trial. The black lines are average trajectories for congruent and incongruent trials.

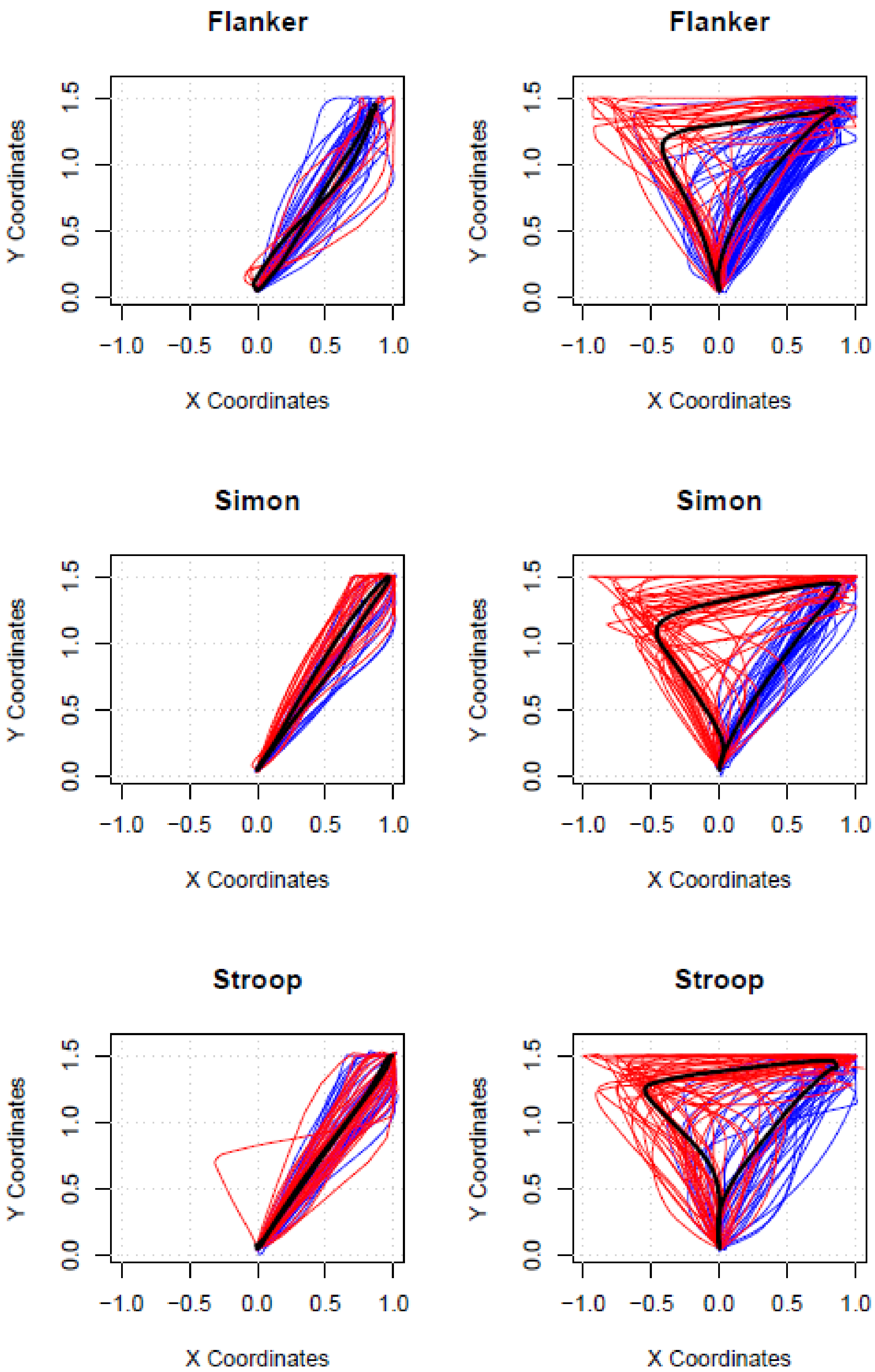


Figure 3.7. Trajectories of two participants on all three tasks separately.

3.4.4 Summary

Given the analysis of individual variability on mouse movement responses, mouse responses were characterised by an unexpectedly high level of variability at the level of individual participants (e.g. inter-individual variability, intra-individual variability). Such variability might be typically obscured in experiments of this type, as most (or all) mousetracking studies would not investigate individual variability. It is unknown whether this is a property typically emerged with the mousetracking paradigm or with bilingualism.

Chapter 4 General Discussion

We conducted a study to investigate the possibility of bilingual advantage on cognitive control with three widely used executive function tasks: Flanker, Simon and Spatial Stroop. Three groups of participants (English monolinguals vs. Chinese-English bilinguals vs. Chinese monolinguals) were compared. Unlike those key-presses experiments previously conducted, we applied the mousetracking paradigm again in this study, so participants make responses with mouse rather than keyboard buttons. The software MouseTracker was used to collect data and to compute mouse movement trajectories for analysis. Accuracy, initiation time, response latency, and mouse movement trajectories were computed and examined.

4.1 Response latency and initiation times

Bilingual advantages in cognitive control?

As expected, incongruent trials in all three tasks consistently resulted in longer response latencies than congruent trials. In addition, in comparison of trajectories between conditions, congruent trajectories in all three tasks were more ideal and straighter, whereas incongruent trajectories were more deflected toward the alternative response. These results are consistent with previous studies via mousetracking paradigm (Incera & McLennan, 2016; Damian, Ye, Oh, & Yang, 2018), that there were congruency effects on response latencies and mouse movement trajectories across tasks.

However, the results for response latencies provided no support for bilingual advantage in any cognitive control. Hypothesised by enhanced inhibitory control or enhanced attentional control hypothesis, bilinguals should suffer less from congruency effects on response latency or have an overall faster response latency, respectively. In fact, bilinguals did not show smaller congruency effects or any ‘global advantage’ on response latencies. Bilinguals showed significantly larger response latencies than Chinese monolinguals, indicating an unexpected “bilingual disadvantage”.

Interestingly, the Chinese monolingual group suffered the most from congruency effects in the Flanker task. With the overall response latencies, the Chinese monolingual group showed the smallest average response latencies in the Spatial Stroop task. These results were not consistent across tasks. One possible explanation for this is that the experiment settings (e.g. environment, temperature, instructions) might influence one's responses. English monolinguals and bilinguals had their experiments done in the similar lab settings and under similar temperature, whereas Chinese monolinguals completed their experiments in a completely different setting with another experimenter in China. However, it is possible the environment factor is not sufficiently strong, so it did not induce a significantly constant pattern. Another potential explanation is Chinese monolinguals might read the stimulus of Flanker task in a more logographic way as how they read Chinese characters. It is likely that Chinese monolinguals consider every arrow in the string (e.g. →→→→→) as a character, whereas English monolinguals see the whole string as a single word. A different perceiving way might take Chinese monolinguals longer to distinguish congruent and incongruent trials. However, since Chinese-English bilinguals have had chance in practicing reading English, they might be able to efficiently process the arrow string as the English monolinguals.

Overall, the results of response latencies (and initiation times) would not support the claim of bilingual advantage in inhibitory control or attentional control, because bilinguals did not show an advantage on response latencies and even 'disadvantage' when compared with Chinese monolinguals. Nevertheless, these results were (not surprisingly) consistent with our previous study (Damian, Ye, Oh, & Yang, 2018) that bilinguals and English monolinguals showed comparable response latencies over three tasks. The current results would also support previous 'null effects' on the bilingual advantage with traditional standard (Paap & Greenberg, 2013; Paap et al., 2015, 2016).

Bilinguals as experts?

It has been discussed in previous studies that whether bilinguals behave like experts.

As experts, bilinguals would initiate their responses later but then make a more efficient response (faster and more idealised trajectories) to the correct answer, regardless of condition. Both Incera and McLennan (2016) and Damian, Ye, Oh, and Yang (2018) found this pattern of group effects on initiation times. However, the current study showed no significant difference in initiation times among groups: the three groups started their movements at approximately the same time.

Potential factors that might result in these inconsistent results have been considered. Incera & McLennan (2016) and the current study display the target stimuli immediately after clicking on the “start” button, while Damian et al. (2018) had a fixation cross before the onset of stimuli display. Some might argue that the removal of the fixation cross is crucial to initiation time, but this is implausible because our experiment structure is very similar to Incera and McLennan (2016)’s with no fixation cross. Another possible factor is that Incera and McLennan would warn their participants if they started the movement later than 500ms. Such warning might have different effects on bilinguals and monolinguals. However, in the current results, it was rare to see responses that were initiated later than 500ms. Therefore, it is still unknown why these inconsistent results emerged.

To summarise these findings: generally, individuals started their movements at around the same time and also arrived at the correct response button the same time among groups. This is perfectly consistent with the null hypothesis that monolinguals and bilinguals do not differ at least in traditional measures (i.e. response latency). However, the current results did not replicate what previous studies found as “expert-like behaviour” in bilinguals.

4.2 Analysis of mouse movement trajectories

Although the results of response latency were expected, the mouse movement trajectories were completely surprising. According to our previous study, bilinguals carried out a more ideal trajectory across conditions (congruent and incongruent) in all three tasks (Flanker, Simon and Spatial Stroop), and we interpreted that overall pattern as enhanced attentional control (group effects on MD and AUC).

However, this pattern was not replicated in the current results and even a pattern in the opposite direction was found. Mouse movements in English monolinguals suffer less from congruency effects compared to the other two groups (in Flanker and Stroop tasks); English monolinguals also generally carried out more efficient mouse movements for both conditions in all tasks, as reflected as group effects on MD and AUC (i.e. significantly smaller MD and AUC for English monolinguals mouse movements in both groups).

Overall, English monolinguals showed the most ideal trajectories in all three tasks for both conditions. Bilinguals consistently showed the most deviation in their trajectories. The (quasi-)Chinese monolinguals generated trajectories somewhat between the other two groups, which makes sense because they are not pure monolinguals but low-proficiency bilinguals. The results of MD and AUC also confirmed this pattern that English monolinguals carried out more efficient responses compared with bilinguals.

4.3 Individual variability

The mouse movement findings are puzzling and provide no support for the BA hypothesis. Moreover, these findings even support "bilingual disadvantage" which has been rarely found in previous studies. Therefore, a deeper analysis of the dataset was computed. First, we plotted average trajectories for every individual (i.e. individual average trajectories; every trajectory represents one individual), regardless of task and condition. The individual average trajectories showed a huge individual variability within each group, varying from very straight lines to extremely deviated lines. However, since these trajectories averaging across conditions, it is difficult to know if such individual variability comes from the congruent condition or incongruent condition.

Then the next step is to compare individual average trajectories between congruent and incongruent conditions. As expected, all individuals showed very ideal average responses in congruent condition. However, individual variability emerges in incongruent condition, varying from very ideal ones to very deviated ones. This is obscured in the overall picture, as congruency effects are evident at the group level across tasks: some individuals are able to make efficient responses for **both** conditions while others show congruency effects between conditions.

Then we mapped individual trajectories separately with their raw trial-by-trial trajectories, investigating whether there is variability within individuals. For the results, it is striking to see variously different shapes of mouse movement patterns. Moreover, even within each individual, the variability is not eliminated, as participants responded to the stimuli very differently trial by trial. That is, a specific individual would not adopt one particular strategy to accomplish their tasks over time. Based on these findings, it warrants further research in individual variability on EF tasks via mousetracking before investigating any bilingual effect on these tasks. Unfortunately, there was no mousetracking study looking into individual variability on similar tasks, so it is unknown if the current findings are a general problem or simply a single case. It may be helpful to examine this issue within monolinguals first and then to compare between monolinguals and bilinguals. One potential factor is due to the uniqueness of mousetracking paradigm, as this technique provides an opportunity for individuals to change their decisions within a trial, which could increase individual variability.

Difference between monolingual groups?

Another important question needs to be answered is whether the two monolingual groups differ in mouse movements, which is the purpose of adding a third group in the current study. In the previous study, we recruited English monolinguals and Chinese-English bilinguals, which might lead to a discussion of potential culture effects rather than bilingualism on EF tasks. English monolinguals and Chinese monolinguals showed similar pattern in response latencies and mouse movement trajectories on Simon and Stroop tasks. No significant difference was found between these two groups on Simon and Stroop. However, in the Flanker task, Chinese monolinguals showed the largest congruency effects on response latencies, and English monolinguals had significantly better mouse trajectories for both conditions. This is not entirely unexpected because in our previous study, the results obtained from the Flanker task also showed somewhat different patterns among groups. Additionally, with a naked eye, the trajectories of Flanker task were not that similar to those of Simon and Stroop tasks, even though they are in the same direction (specificity of the Flanker will be discussed more in limitations). In general, two monolingual groups did not significantly differ in Simon and Stroop tasks, so future research might consider recruiting more than one type of bilinguals as participants.

4.4 Limitations

One obvious limitation of this study is the small sample size, which is inevitable because of limited time and expenses. Given that the effects anticipated from Damian et al. (2018) were not found, the issue arises whether the experiment reported here had appropriate statistical power (e.g., Bakker, 2015). This is a difficult question to answer, as results derived from mouse tracking data are multi-dimensional (e.g., Damian et al. reported five different dependent measures) and from the previous work we have only an incomplete idea of the expected effect sizes on each. If we mainly focus on the Simon task in which Damian et al. found group differences on initiation times which were very pronounced, the bilingual advantage in their findings had a large effect size ($d = .80$). If we anticipate a similar effect size for the current study, a power analysis would suggest $N = 26$ in each group to ensure a power of .80. The sample size of the current study (79 English monolinguals; 34 Chinese monolinguals; 26 Chinese-English bilinguals) might hence be sufficient to capture such large effects, but it would not be large enough if group effects were genuine but medium or small. The sample size of the current study was constrained by practical aspects of participant recruitment;

however, it is acknowledged that future students should aim at recruiting substantially larger samples to ensure adequate statistical power.

We also could not recruit pure Chinese monolinguals, as university students from China all have learned English and passed English tests in high school (as a requirement of university admission). However, they should not be at the same level of English as international students in UK university, because international students are immersed in an English-speaking country and are required to practice/use English in their daily life. International students also switch more often, because they will communicate with friends or families in Chinese but with teachers or classmates in English. Such intense language experience should contribute differences in proficiency and use frequency of English between “Chinese-English bilinguals” and “Chinese monolinguals”. From the results, we can see Chinese monolinguals behave more similarly to English monolinguals for their mouse movements, especially in the Simon and Stroop tasks. Again, results from the Flanker indicate differences in mouse movements among the three groups, particularly for the incongruent condition.

Why is the Flanker task special? Even if we deleted all fixation cross across tasks to make their task design more similar to each other, the trial structure among tasks is still slightly different. Now in three tasks, target display appeared immediately after the click on “start” button. However, participants have 1,700ms to respond in the Flanker task, whereas they must respond within 1,000ms in the Simon and Stroop tasks. The Flanker task also contains 75% of congruent trials and 25% of incongruent trials to increase task difficulty and “monitoring demand”, whereas the other two tasks consist of 50/50 of congruent and incongruent trials. These task designs were taken from Zhou and Krott (2016a) for direct replication. We predicted in our previous study that removing the ‘pre-target period’ should make the three tasks resemble each other. However, this seems not enough, so the next step could be to control their maximum response latency and the percentage of congruency. In the current study, we found most participants could complete each trial for the Flanker task within 1,000ms, so it is reasonable to cut down their maximum response latency to 1,000ms. Because Incera and McLennan (2018) found no bilingualism effect on the Flanker task with 50/50 congruency, future research might consider increase task difficulty for the other two tasks. In addition, as mentioned above, removing the ‘pre-target period’ is probably not a good idea and it might not contribute to the specificity of the Flanker task. The aim of removing fixation cross is to avoid negative initiation time, which is unlikely induced by target stimuli. However, another problem has arisen that the current setting might not best simulate a real daily situation that usually individuals have a preparation time. Therefore, future research needs to determine whether such a pre-target period should be added with a same trial structure for all three tasks.

The measure of monitoring ability/attentional control in this current study is to calculate the mean response latency with mixed trials by group. The rationale is that faster RT represents better monitoring ability. However, global response latency might not be a pure measure of monitoring (Paap et al., 2016). Paap and colleagues proposed that an overall RT advantage might be due to better perceptual-motor skills rather than superior monitoring ability. To ‘purify’ the results, they suggested comparing the mean RT from a pure block of neutral trials to mean RT on the congruent trials in a block with mixed trials (congruent and incongruent). Therefore, the difference between these two mean RTs will filter out perceptual-motor skills or other low-level processing. In addition, the ‘purified’ results should represent the difference between monitoring ability in an interference task and monitoring ability in a control task, whereas the original version cannot reveal the baseline monitoring ability in an interference task that constantly requires conflict detection.

In both our studies, MD and AUC were both reported as measures of mouse movement curvature. However, they are not independent to each other and highly correlated. Positive effects on both MD and AUC should not be considered as double confidence. Based on Freeman and Ambady (2010, p.230), AUC can report the overall attraction towards the unselected response, however, MD might locate the maximum attraction to specific time steps, which might be helpful for exploring individual variability in mouse movements. X-flips, which capture fluctuating movements along the horizontal axis, were not reported in this study. However, it might be useful for investigating individual variability, as it might reveal participants' fluctuations in mouse trajectories while MD and AUC could not show this.

4.5 Future Directions

Explore the individual variability

The individual variability we found in the current results is highly unexpected. For future research, it might be a good idea to explore why such variability is present. Therefore, we could (and will) speculate this issue in detail before we chase down the subtle potential (language) group differences. For example, we could map out whether individuals respond to instructions differently, whether imposition of a response deadline could push every individual to make similarly "early" responses, and/or whether individuals would respond more conservatively to penalties for errors, etc. It is likely that these factors matter more in the BA studies which did between-subject manipulations than those that had done within-subject manipulations.

It is also of speculation that individuals seem to not only apply one decision-making strategy – or there are individual differences in their decision-making strategies. As discussed in Chapter 2, the embodied choice model has proposed that decision and action are inseparable, and action would bias the decision-making process. However, in our individual data, for example, participant 7 in Figure 3.6, would wait for a relatively long time until a decision has been made, and then make an almost straight response. Therefore, "dynamic action" might not be always present, and decision making and response execution might be flexible. It is possible that these strategies are not discrete but on a continuum.

It will also be imperative to assess individual differences in response trajectories via inferential statistics, rather than just descriptively (as in, e.g., Figures 3.6 and 3.7).

Conventional inferential statistics which are based on averages by participants and conditions are not well-suited for the exploration of individual differences. However, recently developed analyses based on “linear mixed effect” models might provide further insight. In these analyses (e.g., Baayen, Davidson & Bates, 2008), statistical analysis is based on raw data rather than averages, and individual differences can be explicitly incorporated into model equations. Some templates exist as to how individual differences can be statistically captured in such models (see, e.g., Kliegl, Wei, Dambacher, Yan & Zhou, 2011). The next step in my pursuit of the bilingual advantage as revealed in mouse tracking will consist of stepping back from the issue of linguality, and first gaining a conceptual and statistical handle on how individuals make decisions, and carry out responses, in mouse tracking tasks.

Switch task

The current study is mainly about inhibitory control and attentional control (monitoring). However, another aspect of cognitive control which is currently understudied and in which a BA could emerge, is switching. Switching is important for bilinguals because bilinguals constantly switch between languages. Miyake and colleagues (2000; 2012) considered switching as one component of executive function, which provides a theoretical basis for potential enhanced switching in bilinguals. Hence, the question is: Does the underlying mechanism of switching between languages the same as that in non-verbal task switching?

Positive findings have been reported to support bilingual advantages in enhanced switching. For example, Prior and MacWhinney (2010) found bilingual advantages in cued-switching tasks, and Prior and Gollan (2011) replicated this finding. Measures of switching also typically show higher test-retest reliability and convergent validity, indicating those measures as appropriate ways of investigating switching component. Most studies on switching usually asked participants to switch between tasks during their experiments, such as the colour-shape switching task. The target stimuli usually are colour (red/blue) and shapes (circle/triangle), and there are two aspects of every stimulus (colour and shape). In a classic colour-shape switching task, participants are required to generate a response under specific task instructions depending on the cues. For example, if Cue 1 is displayed before a stimulus, then participants should judge the colour of the stimulus and press the corresponding key (e.g. red for left, blue for right); if Cue 2 is displayed before a stimulus, then participants should judge the shape of the stimulus and press the corresponding key (e.g. triangle for left, circle for right). Therefore, as Cue 1 and Cue 2 are randomised and mixed in one experimental block, participants have to switch back and forth between colour and shape tasks.

However, a review of 17 studies and 53 tests in this literature by Paap and colleagues (2017) pointed out that except the two studies we mentioned above (Prior & MacWhinney, 2010; Prior & Gollan; 2011), others all failed to replicate the positive findings. Paap et al. (2016) also could not observe any enhanced switching ability in bilinguals in the subsequent research. However, all those studies were done with key presses, and mousetracking might be a more sensitive tool to detect any BA in task switching. Corresponding experiments are conducting with a colour-shape switch tasks in mousetracking version.

Dual mechanisms of control framework and AX-CPT

Braver (2012) suggested an emphasis from the diversity of cognitive control functions to its intrinsic variability in terms of “temporal dynamics”. He described the dual mechanisms of control (DMC) framework, which proposes that cognitive control may include two distinct modes: proactive control and reactive control. Proactive control is a mode that actively maintains or sustains goal-relevant information, in preparation for incoming cognitive tasks. This description sounds similar to the “monitoring abilities” or “attentional control”. Reactive control is a mode that detects and corrects any interference after stimuli onset. To be more specific, in a classic Stroop task, participants need to judge the colour of a word (e.g. BLUE in blue or BLUE in green) but ignore the text itself. Based on this framework, proactive control would keep the task instructions online and help individuals prepare for it before stimuli onset, whereas the reactive control will correct the response if the participant is biased by the word semantics or reactivate the task instructions after the stimuli onset. This framework did not consider cognitive control as independent components but a mechanism with two modes that are weighted differently. Therefore, individuals could have a preference for one of two modes while the most important thing is to adjust them in a more efficient way. From this framework, cognitive control is combination of two modes rather than discrete components (i.e. monitoring and inhibitory control). To be noted, these two modes are not necessarily alternative, and cognitive control is the dynamic combination of these two modes.

This framework provides a new perspective of investigating cognitive control. DMC framework also indicated the possibility of intra-individual variation and inter-individual variation. Braver argued that the interference expectancy might change one's weighing between proactive and reactive control strategies. For example, in the low expectancy condition, reactive control is recruited, whereas proactive control will be preferred in the high expectancy condition. In addition, in the high cognitively demanding condition, reactive control is more observed, but in the low cognitively demanding condition, the proactive pattern is more observed. According to these findings, the Flanker task used in the current study might not be appropriate, as it might have biased participants' cognitive control strategy by enhanced task demanding. He also highlighted that personality trait such as anxiety and trait reward sensitivity could be associated with inter-individual variation.

In terms of between-group variation, the AX-CPT (AX-Continuous Performance Task) has become a popular paradigm to target differences in the use of two cognitive control strategies. For example, in a version of this task (adapted from Ophor et al., 2009, cited in Morales et al., 2013), stimuli were cue-probe pairs, such as AX with A as a cue and X and a probe. Participants were required to respond "yes" to the target cue-probe combination (AX), and to respond "no" to other cue-probe combinations (e.g. BY, BX, AY). Task demands can be manipulated by inserting distractors between every cue and probe (e.g. "yes" to A-FMD-X). To achieve better performance and fewer errors, it requires efficient adjustments of both monitoring abilities and inhibitory control in a temporally dynamic way, because different conditions require a different degree of engagement of proactive control and reactive control, while our cognitive resources are limited.

Morales and colleagues (Morales et al., 2013) compared monolinguals to bilinguals in this AX-CPT paradigm. They found that bilinguals outperformed monolinguals in AY condition. They suggested that all individuals adopt monitoring to facilitate their performance in target condition (AX) by producing target expectancies after an A-cue. However, this high expectancy might not be helpful in AY condition. Thus, the optimum performance in AY condition requires "higher adjustment of proactive and reactive control" (Morales et al., 2013), to induce proper suppression of the "yes" tendency. Then they interpreted that bilinguals can better adjust the proactive-reactive control dual mechanism. They also pointed out that this better regulation of two modes might explain why bilinguals have longer response latency in their study, as they took time to shift to the reactive control and engaged in it more than monolinguals. Their ERP experiment (Morales et al., 2015) also suggested an effect of bilingualism on proactive and reactive control. According to this study, proactive control is reflected as P3b, which is elicited during cue-probe interval; reactive control is indicated by N2 and P3a, which are observed after probe. Their results replicated their previous behavioural findings and showed that ERP between monolinguals and bilinguals differed in reactive control-related components (N2 and P3a). Bilinguals showed more negative N2 and higher positive mean P3a amplitudes to AY conditions than monolinguals, which are thought to reveal better reactive control. However, such group difference did not show on BX and BY trials. Therefore, they indicated that bilinguals showed better reactive control than monolinguals only when competing cue information exists.

Overall, their studies might provide some explanations about the inconsistent findings in this literature and a new perspective for further research. Replication of these results with mousetracking paradigm might provide a greater amount of dynamic information between bilinguals and monolinguals.

4.6 Conclusion

In conclusion, the present study did not find 'expert-like behaviour' in bilinguals. Chinese-English bilinguals had comparable initiation times and response latencies compared with English monolinguals. The present results also found different mouse movement patterns among groups, indicating inconsistent results with previous studies. English monolingual group showed more ideal trajectories in all three tasks and Chinese monolingual group showed a similar pattern except in the Flanker task to English monolingual group. In contrast, bilinguals indicated more deflection in their mouse movements towards alternative responses in both congruent and incongruent conditions, especially in the Flanker task. These results generally revealed no bilingual advantage in these EF tasks or even 'bilingual disadvantage'. These findings encouraged us to examine the data more deeply to the individual level. Surprisingly, unexpectedly large individual variability was found in each group. Extremely different mouse trajectories could appear within a group and within an individual. Nevertheless, given limited empirical evidence, it is difficult to know whether such large individual variability is common in mousetracking experiments on bilingualism research, and where this variability comes from. One potential explanation might be that individuals adopted various strategies (proactive control vs. reactive control) in the tasks to resolve conflicts. However, further research is necessary to investigate whether this is the case, for example, via AX-CPT. Although this study did not replicate previous findings, the mousetracking paradigm is still a sensitive technique that has the potential to observe cognitive processing that is obscured by button presses experiments. Therefore, it is recommended to replicate other previous studies with the mousetracking paradigm to see if there is any bilingualism effect obscured by traditional measures.

References

Abutalebi, J., Canini, M., Della Rosa, P. A., Green, D. W., & Weekes, B. S. (2015). The neuroprotective effects of bilingualism upon the inferior parietal lobule: a structural neuroimaging study in aging Chinese bilinguals. *Journal of Neurolinguistics*, 33, 3-13.

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., ... & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral cortex*, *22*, 2076-2086.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*, 207-245.

Bakker, M. (2015). Power problems: n> 138. *Cortex; a journal devoted to the study of the nervous system and behavior*, *73*, 367.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*, 160-166.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *60*(1), 68.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism*.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological bulletin*, *143*, 233.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, *11*, 485-508.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, *19*, 290.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, memory, and cognition*, *34*, 859.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*, 103-119.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in cognitive sciences*, *16*, 106-113.

Buchsbaum, B. R., Olsen, R. K., Koch, P., & Berman, K. F. (2005). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron*, *48*, 687-697.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*, 135-149.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*.

Damian, M., Ye, W., Oh, M., & Yang, S. (Accepted/In press). Bilinguals as "experts"? Comparing performance of mono- to bilingual individuals via a mousetracking paradigm. *Bilingualism: Language and Cognition.*

De Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias?. *Psychological science*, *26*, 99-107.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*, 19-23.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*, 143-149.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*, 226-241.

Freeman, P. W., & Lemen, C. A. (2008). A simple morphological predictor of bite force in rodents. *Journal of Zoology*, *275*, 418-422.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 640.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, *16*, 562.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515-530.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review*, *18*, 625-658.

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: A multimodal imaging study. *Journal of Neuroscience*, *33*, 13663-13672.

Incera, S., & McLennan, C. T. (2016). Mouse tracking reveals that bilinguals behave like experts. *Bilingualism: Language and Cognition*, *19*, 610-620.

Incera, S., & McLennan, C. T. (2018). Bilingualism and age are continuous variables that influence executive function. *Aging, Neuropsyc*

Klein, D., Mok, K., Chen, J. K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: a cortical thickness study of bilingual and monolingual individuals. *Brain and language*, *131*, 20-24.

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology, 1*, 238.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, pnas-0811323106.

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., & Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*.

Lepora, N. F., & Pezzulo, G. (2015). Embodied choice: how action influences perceptual decision making. *PLoS computational biology*, *11*, e1004110.

Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, *31*, 16808-16813.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: language and cognition*, *11*, 81-93.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current directions in psychological science*, *21*, 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*, 49-100.

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology*, *25*, 531-546.

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18). Springer, Boston, MA.

Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children* (Vol. 2). Multilingual Matters. *Proceedings of the National Academy of Sciences*.

Olsen, R. K., Pangelinan, M. M., Bogulski, C., Chakravarty, M. M., Luk, G., Grady, C. L., & Bialystok, E. (2015). The effect of lifelong bilingualism on regional grey and white matter volume. *Brain research*, *1612*, 128-139.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*, 15583-15587.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*, 232-258

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278.

Paap, K. R., Johnson, H. A., & Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue?

Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive*

Psychology, *29*, 89-112.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annu Rev Neurosci*.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human brain mapping*, *19*, 170-182.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*, 415-422.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *URL https://www.R-project.org/.*

Appendix A

Below is the Language History Questionnaire adapted from Silverberg and Samuel (2004).

Language History Questionnaire

Part. No. ______

Please indicate your self-perceived proficiency in **English** by drawing a **vertical line** on the scale below. The far left end stands for no knowledge in English, and the far right end stands for 100% native-like proficiency.

	Little/No knowledge		Native-like
Speaking	0	|---------------|	100
Understanding speech	0	|---------------|	100
Reading	0	|---------------|	100
writing	0	|---------------|	100

1. Please indicate your self-perceived proficiency in **Chinese** by drawing a **vertical line** on the scale below. The far left end stands for no knowledge in Chinese, and the far right end stands for 100% native-like proficiency.

	Little/No knowledge		Native-like
Speaking	0	|---------------|	100
Understanding speech	0	|---------------|	100
Reading	0	|---------------|	100
writing	0	|---------------|	100

2. Please draw a vertical line on the scale to indicate the current use of both **English and Chinese** in oral communications, at home and outside home.

	No English		All English
At home	0	|---------------|	100
Outside home	0	|---------------|	100
	No Chinese		**All Chinese**
At home	0	|---------------|	100
Outside home	0	|---------------|	100

3. Do you speak Chinese (Mandarin or Cantonese) fluently? (By fluently we mean that, for everyday conversations, you are able to converse with native speakers without having to consciously translate).

Yes No

4. Please list all languages you speak (which reached **native or native-like competence**) in the order you began to acquire them (since born). Indicate at what age you began to learn each and at what age (approximately) you mastered each:

	Language	Age began to learn	Age mastered
1			
2			
3			
4			

5. In what setting did you acquire your second (and third, if applicable) language? (e.g., at home, through school, living abroad, other)

Second language	**Third language**	**Forth language**
o At home	o At home	o At home
o Through school	o Through school	o Through school
o Living abroad	o Living abroad	o Living abroad
o Other (please specify) ____________	o Other (please specify) ____________	o Other (please specify) ____________

6. **Language(s) of parents** (or primary caretaker, guardian, etc):________________

7. Please roughly describe your previous language use history using the table provided.

Please specify your own language use experience at different stages of your life till now.

Age :			
Only use **first** language regularly			
Only use **second** language regularly			
Use both languages regularly, but in **different** settings (e.g., use Chinese at home and English outside of home)			
Use both languages regularly, but in **the same** setting (e.g., use both languages both at home and outside of home)			

8. **Current language use** (check the one that applies)

Do you now:

_______use primarily one language? If so, which one? _________

_______use both languages regularly but in different settings (i.e., one at home and one at school, one with friends and one with family, etc.)

_______use both languages every day within the same setting (i.e., use both at home)

9. Do you have friends or family who are also bilingual in the two languages you speak?

Yes No

10. When speaking with these bilingual friends/family members, do you ever find yourself using both languages within the same conversation or even in the same sentence?

____________ Yes, frequently

____________ Yes, but only rarely

____________ No, never

11. **True/False:**

T F I mix languages only when talking to friends or family.

T F I mix languages in conversations with other bilinguals because this enables me to express myself better.

T F I mix languages because of other reasons (please specify): ___________

T F I try not to mix languages in the same conversation.

12. What is the highest level of certificate that you have got now?

1. GCSE or below		6. Others (please specify)	
2. A level			
3. Bachelor's degree			
4. Master's degree			
5. Doctorate or above			

13. If one of your languages is Chinese, please indicate which dialect you speak (e.g., Mandarin, Cantonese etc . . .)

____